April 1987 UILU-ENG-87-2226

COORDINATED SCIENCE LABORATORY
College of Engineering

REPRESENTATION AND THREE-DIMENSIONAL INTERPRETATION OF IMAGE TEXTURE: AN INTEGRATED APPROACH

Dorothea Blostein
Narendra Ahuja

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS None
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; distribution unlimited
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) UILU-ENG-87-2226	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Coordinated Science Lab University of Illinois 6b. OFFICE SYMBOL (If applicable) N/A	7a. NAME OF MONITORING ORGANIZATION Air Force Office of Scientific Research Eastman Kodak
6c. ADDRESS (City, State, and ZIP Code) 1101 W. Springfield Avenue Urbana, IL 61801	7b. ADDRESS (City, State, and ZIP Code) Bolling AFB, Washington, D.C. 20332 Corporate Headquarters, Rochester, N.Y. 14692
8a. NAME OF FUNDING / SPONSORING ORGANIZATION Air Force Office of Scientific Research and Eastman Kodak 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER AFOSR 87-0100
8c. ADDRESS (City, State, and ZIP Code) Bolling AFB, Washington, D.C. 20332 Corporate Headquarters, Rochester, N.Y. 14692	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)
Representation and Three-dimensional Interpretation of Image Texture: An Integrated Approach

12. PERSONAL AUTHOR(S)
Blostein, Dorothea and Ahuja, Narendra

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Technical	FROM _____ TO _____	April 1987	130

16. SUPPLEMENTARY NOTATION

17. COSATI CODES		
FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
texture, texels, texture gradients, surface shape, homogeneous image, planar surfaces

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

A perspective view of a slanted textured surface shows systematic changes in the density, area and aspect-ratio of texture elements. These apparent changes in texture element properties can be analyzed to recover information about the physical layout of the scene. However, in practice it is difficult to identify texture elements, especially in images where the texture elements are partially occluded or are themselves textured at a finer scale. To solve this problem, it is necessary to integrate the extraction of texture elements with the recognition of scene layout. This paper presents a method for recovering the orientation of textured surfaces while simultaneously identifying texture elements. Candidate texture elements are constructed from overlapping circular regions of relatively uniform gray-level. The uniform circular regions are found by convolving the image with $\nabla^2 G$ (Laplacian-of-Gaussian) masks over a range of scales, and comparing the convolution output to that expected for a circular disk of constant gray level. True texture elements are selected from the set of candidate texture elements by finding the planar surface that best predicts the properties of the candidate texture elements. A planar fit is evaluated by comparing the predicted texture-element areas to the actual areas of the candidate texture elements. The planar fit receiving support from the most regions is chosen as the correct interpretation. Simultaneously, those candidate texture elements that support the best plane are identified as the true texture elements. Results are shown on images of many natural textures, including rocks, leaves, waves, flowers, bark, and clouds. Textures consist of both bright and dark regions, corresponding to lit and shadowed areas, or to foreground and background. The positive-contrast and negative-contrast regions of each image are analyzed separately. For a number of images used in our experiments, the two analyses result in slant and tilt estimates that are within ten degrees of each other. For other images, the discrepancy is larger because of implementation restrictions or because these textures violate the homogeneity assumptions made in one or both of the analyses.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code) / 22c. OFFICE SYMBOL

ABSTRACT

A perspective view of a slanted textured surface shows systematic changes in the density, area and aspect-ratio of texture elements. These apparent changes in texture element properties can be analyzed to recover information about the physical layout of the scene. However, in practice it is difficult to identify texture elements, especially in images where the texture elements are partially occluded or are themselves textured at a finer scale. To solve this problem, it is necessary to integrate the extraction of texture elements with the recognition of scene layout. This paper presents a method for recovering the orientation of textured surfaces while simultaneously identifying texture elements. Candidate texture elements are constructed from overlapping circular regions of relatively uniform gray-level. The uniform circular regions are found by convolving the image with $\nabla^2 G$ (Laplacian-of-Gaussian) masks over a range of scales, and comparing the convolution output to that expected for a circular disk of constant gray level. True texture elements are selected from the set of candidate texture elements by finding the planar surface that best predicts the properties of the candidate texture elements. A planar fit is evaluated by comparing the predicted texture-element areas to the actual areas of the candidate texture elements. The planar fit receiving support from the most regions is chosen as the correct interpretation. Simultaneously, those candidate texture elements that support the best plane are identified as the true texture elements. Results are shown on images of many natural textures, including rocks, leaves, waves, flowers, bark, and clouds. Textures consist of both bright and dark regions, corresponding to lit and shadowed areas, or to foreground and background. The positive-contrast and negative-contrast regions of each image are analyzed separately. For a number of images used in our experiments, the two analyses result in slant and tilt estimates that are within ten degrees of each other. For other images, the discrepancy is larger because of implementation restrictions or because these textures violate the homogeneity assumptions made in one or both of the analyses.

The support of the Air Force Office of Scientific Research under grant AFOSR 86-0009 and the Eastman Kodak Company is gratefully acknowledged.

CONTENTS

1. INTRODUCTION

Texture variations provide important cues for recovering the three dimensional structure of the surfaces visible in an image. A uniformly textured surface undergoes two types of distortions during the imaging process. Firstly, an increase in the distance from the surface to the viewer causes a uniform compression of increasingly large areas of surface onto a fixed area of image. Secondly, as the surface slants away from the image plane foreshortening causes an anisotropic compression of the texture. The resulting texture gradients provide information about the relative distances and orientations of the textured surfaces visible in an image. Such shape information may be extracted from a textured image independently of texture recognition and classification processes. This paper investigates methods for computer-based extraction of the spatial layout of textured surfaces visible in an image.

1.1. Texture

Texture is an elusive concept, difficult to define precisely. Muerle [1970, page 371] states that

> ...we meet the first problem in using a computer for extracting information about visual texture from a picture - a precise definition of texture does not exist.

and goes on to say that

> ...the primary attributes of a visual texture are *many* variations and *repetitive* variations.

For our purposes, we define texture as the visible variation within an area perceived as a single region. Two points are noteworthy: firstly, texture is a property of a surface, and secondly, texture perception depends on scale. For example, imagine sitting in a packed stadium watching a football game. Looking at the spectators across the field, you see a crowd texture in which each spectator is a texture element. This texture is perceived as a surface. Scale is critical in this perception: looking at the spectators sitting next to you, you do not perceive them as texture elements, nor do you consider yourself as part of a surface. The physical structure of the world is hierarchical; large objects are perceived as structure, and the little sub-objects of which they are composed are texture. As a texture element is approached it resolves into an object that is itself textured.

1.2. Texels

The term *texel*, short for *texture element*, denotes the repetitive unit of which a texture is composed. "Texel" refers to the physical texture element in the real world as well as to the appearance of the texture element in the image. In cases where the distinction must be made, we use the phrases *physical texel* versus *image texel*. Distance and foreshortening changes alter the appearance of the image texel, although the physical texel remains unchanged.

Textures vary in how clearly delineated their texels are. Textures composed of separate physical entities have clearly identifiable texels: each rock in Figure 5(a) is a texel, each house in Figure 7(a) is a texel. Other textures, such as the tree-bark of Figure 19(a) or the waves of Figure 33(a) consist of texels that are less clearly defined. In these textures the perceived location of texel-boundaries may vary slightly from viewer to viewer.

We restrict image texels to be regions of relatively uniform gray level. Under this definition, a physical texel can give rise to several image texels: typically the physical repetitive unit of a texture contains

both bright and dark regions. As described below, we treat the bright and dark image texels as separate texture fields. Requiring an image texel to have "relatively uniform" gray-level means that the texel is uniform relative to the gray-level changes that occur at its own scale; however, the texel may contain significant internal variations of gray level. In other words, large texels appear as regions of uniform gray-level only after suitable blurring of the original image.

1.3. Texture gradients

The term *texture gradient*, in use since Gibson [1950], denotes the systematic texture changes visible across the perspective view of a textured surface. A variety of texture gradients may be defined, depending on which attribute of texture is considered -- there are gradients of apparent texel size, apparent texel density and apparent texel shape. Texture gradients are discussed in detail in Section 2.

1.4. Texture fields

We use the term *texture field* or *field of texels* to denote a collection of image texels that exhibit one or more consistent texture gradients. Consistency is defined with respect to the texture gradients expected from a particular surface arrangement viewed under perspective. There are several common reasons for separate texture fields to occur in a single image. Firstly, many textures are composed of closely associated bright and dark fields which arise from lighting effects. For example, the aerial view of houses in Figure 7(a) contains a field of bright texels composed of the houses and a field of dark texels composed of the shadows cast by the houses. Secondly, associated bright and dark texture fields can arise from the physical structure of the texture elements; see, for example, the sunflowers in Figure 17(a). Thirdly, it is possible for physically separated textured surfaces to be spatially interleaved in an image. This is strikingly illustrated by the birds over water shown in Figure 9(a), where the birds and the water occur in two physically separated planes. Finally, multiple texture fields result from physical surfaces that are covered by several types of texture elements. An aerial view of a residential neighborhood shows one texture field consisting of houses and another texture field consisting of trees.

The concept of texture field is useful for separating portions of physical texels that exhibit differing foreshortening properties. Consider, for example, an aerial view of many flat-roofed houses. The roofs of the houses, which are parallel to the textured plane, are foreshortened increasingly as the angle between the line of sight and the plane decreases, whereas the walls of the houses exhibit the opposite behavior since they are perpendicular to the textured plane. Any analysis of foreshortening in such an image must treat these two texture fields separately. The difference in gray-level properties of the two fields can help to achieve this separation.

1.5. Slant/tilt encoding of surface orientation

From a viewer's perspective, a surface can be represented by specifying the distance to each point on the surface and the unit surface normal at that point. The two degrees of freedom needed to specify a surface orientation can be encoded in a variety of ways. Stevens [1983a] and [1983b] presents arguments in favor of a slant/tilt encoding. Slant and tilt express the orientation of a planar surface relative to the image plane. Slant is the angle between the surface and the image plane. If the slant is zero the surface is parallel to the image plane; we call this a *frontal view* of the surface. On the other hand, if the slant is large the surface recedes steeply away from the viewer. Slant ranges from 0° to 90°. Tilt is the direction in which the surface normal projects in the image; thus the tilt is the direction in the image in which the surface distance increases the fastest. Tilt ranges from 0° to 360°; a tilt of 0° indicates that distance to the viewed surface increases fastest toward the right side of the image. To illustrate the definition of "slant"

and "tilt", we show synthetic textures at various slants and tilts in Figure 1.

1.6. Scope of this work

This work investigates how to exploit textural cues to infer the relative distance and orientation of the textured surfaces depicted in an image. We do not address the problem of texture discrimination or identification.

A primary goal of this work is to demonstrate the feasibility of extracting useful measures of texture gradients from images of natural (as opposed to man-made) textures. The textures present on man-made objects frequently exhibit regularities such as parallel lines, perpendicular lines, equally-sized texture elements, or equally-spaced texture elements. Several existing shape-from-texture algorithms exploit these regularities (Section 3); however, most naturally occurring textures are too variable to permit successful application of these methods. Our results permit fairly successful analyses of natural textures.

A second goal of this research is to develop a uniform treatment of various texture gradients. As discussed in Section 2, any combination of gradients (systematic changes in texel area, aspect ratio, contrast, density) may be present in an image, and the relative accuracy of the gradients varies from image to image. Therefore, we need a unified method of analyzing the variations in different textural properties, and a way to selectively pay attention to the relevant and accurate gradients. Our work provides a start in this direction, but much remains to be done before this goal is fully realized.

A major challenge in texture analysis is to handle scale consistently. Natural surfaces exhibit a rich hierarchy of textures, with each texture element containing subtextures. All texture measurements are prone to distortion due to the presence of subtexture, since the imaging process captures more subtexture details for close texture elements than for distant ones. The algorithms presented in this paper provide good surface-orientation estimates even in the face of significant sub- and supertexture.

1.7. Overview

The organization of this paper is as follows. In Section 2 we begin with a general discussion of texture gradients. After characterizing frontal views of textures, we describe the texture distortions that arise due to changing foreshortening and changing distance. The computer vision literature relating to surface estimation from texture is reviewed in Section 3.

Section 4 presents one of the central ideas of our work, namely, that the extraction of texture elements is an essential step in texture analysis. Texel identification permits correct analysis of texture gradients in images where the texture elements are themselves textured at a finer scale. We review existing methods for texture analysis, which generally do not involve texel identification. Much previous work has avoided texel identification because of its difficulty. However, no adequate substitutes exist. Texture elements cannot be identified in isolation since texels are defined only by the repetitive nature of the texture as a whole. Therefore the identification of texture elements is best done in parallel with the estimation of the shape of the textured surface. We integrate these two processes by first constructing a large set of candidate texels, and then using a surface-fitting algorithm to identify the true texels while simultaneously constructing an approximation to the shape of the textured surface.

In Section 5 we describe a multi-scale region detector that forms the basis of our texel extraction. The region detector, which has a simple implementation and shows robust performance on a wide variety of images, is used to construct a set of candidate texels.

Section 6 presents an analysis of texture gradients in images of textured planes. This analysis is used in Section 7 to develop an algorithm for finding the best planar fit to the candidate texels, while

simultaneously choosing the true texels from among the candidates.

Section 8 discusses the results of the computer analysis on a variety of texture images. A common complaint about computer vision algorithms is that they are not tested on enough images, so the generality of the method remains in doubt. We use seventeen images of natural textures to illustrate the generality of the method and the strengths and weaknesses of the implementation.

We conclude in Section 9 by summarizing the main ideas of the paper.

2. PROJECTIVE DISTORTION AND TEXTURE GRADIENTS

In this section we discuss the various texture gradients that arise due to the imaging process. These gradients convey information about physical scene layout.

2.1. Regularities in frontal views of textures

It is possible to recognize texture gradients despite the inherent variability of natural textures. This is because textures show statistical regularities in a frontal view (in a *frontal view* the textured plane is parallel to the image plane). These regularities are distorted in a systematic and recognizable way by the imaging process.

What texture features tend to be regular? The literature on texture representations (Section 3.1.) describes various methods of characterizing texture regularities. Texel area often shows statistical regularity: the observed texel areas are distributed randomly around an unchanging mean value. Intrinsic texel properties that may be fairly uniform -- in a frontal view with constant lighting -- include the texel area, shape attributes such as aspect ratio, and intensity attributes such as contrast and mean gray-level. In addition to uniformities of intrinsic texel properties, most textures exhibit some regularity of texel placement or density. Many natural processes result in independently placed texels (leaves falling off of a tree, sand piled on a beach), so that local texel density is distributed randomly around an unchanging mean value. In more constrained textures, such as snake skin or brick walls, texels are arranged with near grid-like regularity.

Some textures are not regular in the ways described above. For example, the texels in a pine cone decrease in area toward the top of the pine cone; thus, the physical texels do not have sizes that are distributed randomly around an unchanging mean value. Textures of this type are not suitable for the analyses described in this paper; given only a single view of a texture, it is impossible to distinguish trends in the physical size of texels from trends that arise due to foreshortening and distance changes. Additional cues, such as shading, might help to make this distinction. This subject is beyond the scope of our research.

2.2. Texture gradients for an idealized texture

Projective distortion affects many texture features. Consider first an idealized texture consisting of nonoverlapping circular disks of constant size, as shown in Figure 1. The disks project as ellipses in the image. The major axis of each ellipse is perpendicular to the tilt, whereas the minor axis is parallel with the tilt. The apparent size of the major axes decreases linearly in the direction of tilt, due to increasing distance from the viewer. The apparent size of the minor axes decreases more rapidly: in addition to the distance scaling, the minor axes are reduced by increasing foreshortening. (Foreshortening is inversely proportional to the cosine of the angle between the line of sight and the surface normal.) These changes in the major and minor axes cause an increase of the eccentricity of the ellipses in the tilt direction. The area of the ellipses decreases fastest in the direction of tilt. This is accompanied by an increase in the density of the ellipses. In this idealized texture, the grid-like layout of the texture elements results in linear perspective cues; however, such regularity in texel spacing is extremely rare in natural textures.

2.3. Texture gradients in natural textures

The changes observed in synthetic textures occur in natural textures as well. However, the texture gradients are not as easily observed because natural textures display considerable variability of texel size, shape and density. Physical texels are typically three-dimensional, in contrast with the two-dimensional disks portrayed in Figure 1. This three-dimensionality results in highlights and shadows, and in occlusions between one texel and the next. Also, physical texels have a complex structure. In contrast to a uniform synthetic disk, a physical texel changes appearance as the resolution is increased: subtexture becomes visible. In an image with fixed resolution, more subtexture is visible for the nearby texels than for the distant texels. Supertexture may be apparent in parts of the image: distant physical texels appear as image texels that are small enough to blur into larger regions of relatively homogeneous gray level. These factors make it difficult to identify texture elements and extract texture gradients from real images.

We have defined a texture field as a collection of image texels that exhibits one or more consistent texture gradients. The statistical nature of texture regularities makes it impossible to judge a priori whether two texture elements belong to the same texture field. The perception of a texture field is an aggregation phenomenon that requires a consistent texture gradient across the whole field.

A given texture may be more regular in some features than in others. Therefore the relative accuracy of the various texture gradients may vary from image to image. This is illustrated by the following examples. It is common for texels to be fairly uniform in size and shape, but for the gaps between the texels to be much less uniform. This is illustrated by the birds in Figure 9, the people in Figure 21, the flowers in Figure 25, and the water lillies in Figure 27. In these images, it is more accurate to infer a three-dimensional surface from the size and aspect-ratio gradients than from the gradient of spacings between texels. Our results reflect this: for the flowers image, the planar fit obtained from the area gradient of texels (positive-contrast regions, Figure 25) is much more accurate than the planar fit obtained from the area gradient of the space between the texels (negative-contrast regions, Figure 26). The potential accuracy of the aspect ratio gradient is higher in textures where the physical texels are separated by gaps than in textures where the physical texels overlap and occlude one another. For example, the lilly pads in Figure 27 show a much better aspect ratio gradient than do the rocks in Figure 5. For the water hyacinths of Figure 31, the random three-dimensional arrangement of the leaves makes the aspect ratio gradient very weak, while the area gradient is still quite usable. In images with partial occlusions, such as the movie audience of Figure 15 and the sunflowers of Figure 17, the perspective gradient (length of the unforeshortened texel dimension) is more accurate than the area gradient: if only part of a texel is occluded, the apparent texel area is decreased, whereas the complete unforeshortened dimension (maximum width in the direction perpendicular to the tilt) may remain in view.

2.4. Psychophysics experiments relating to texture gradients

As we have seen, a variety of texture gradients may be defined, depending on which attribute of texture is considered. Cutting and Millard [1984] discuss, among others, the *size gradient* (texel area), the *perspective gradient* (length of the unforeshortened texel dimension), the *compression gradient* (length of the foreshortened texel dimension), the *aspect ratio gradient* (ratio of foreshortened to unforeshortened texel dimension), and the *density gradient* (number of texels per unit image area). Rosinski and Levine [1976] mention that these gradients are mathematically equivalent in that, if the gradients could be measured with perfect accuracy, each one would provide the same information. However, the gradients vary in their perceptual effectiveness: they are not equivalent in terms of an observer's ability to extract or use them. The psychology literature contains reports of many experiments that address this subject. These experiments provide interesting insights into texture perception; however, since the experiments are

performed on highly idealized synthetic textures, the results may not generalize to textures occurring in nature.

2.4.1. The role of texture cues in perception

Following the pioneering work of Gibson [1950] and [1966], many researchers have studied the roles of various texture cues in surface perception, using experiments with idealized synthetic textures to determine the relative effectiveness of the various texture cues. Some of the early work in psychophysics centers on the relevance of two image properties for judging slants of planar surfaces: (1) the projective distortion in the shape of a single object versus (2) the gradient of object sizes across the visual field (accompanied by a gradient in object density). Flock ([1964], [1965]) emphasizes the role of the size/density gradients, whereas Freeman ([1965], [1966a], [1966b]) argues that the foreshortening of an object's shape is responsible for perception of surface slant. Freeman even suggests that texture gradients have no role to play in surface slant perception by humans: he compares subjects' judgements of surface slant from a textured surface and from a textureless rectangle. These disagreements are due, at least in part, to inappropriate test data and to the ill-defined nature of the problem. Braunstein and Payne [1969] provide further relevant discussion.

Gruber and Clark [1956] focus on the relationship between texture density and slant perception. They use synthetic disk textures to conclude that the impression of slant is maximized at a particular texel density (which varies with texel area); stimuli with a lesser or greater texel density give rise to a weaker slant perception. Eriksson [1964] obtains similar results.

2.4.2. Relative importance of various texture gradients

Many experiments have been performed to test the relative importance of various texture gradients. Braunstein and Payne [1969] use dot and line patterns to conclude that linear perspective appears to be the principle variable underlying relative slant judgements. Phillips [1970] uses disk textures to test the relative importance of size, shape and density information, but warns that it would be improper to generalize his results to other types of visual texture. Phillips finds in his experiments that slant judgements depend less on texel density than on texel size and shape parameters (texel attributes that could be responsible for the slant judgements include texel area, aspect ratio, major axis length and minor axis length). Rosinski and Levine [1976] find that minor axis length is a less effective cue than major axis length or texel area. Attneave and Olson [1966] experiment with grid and line textures to test the relative importance of contour-density and texel-size cues, but their measures are so specific to their test patterns that the results are difficult to generalize. Several different properties of image texture that capture surface information, and the effectiveness of these properties in human vision, are reviewed by Rosinski [1974].

Vickers [1971] was among the first to advocate an approach involving accumulation of evidence from multiple texture gradients. Vickers' principle of perceptual economy states that the magnitude and strength of slant judgment are related to the amount of total evidence present in favor of the judgment. Support for this principle comes from experiments that show that increasing the number of texture gradients causes a more vivid tridimensional impression, increases the judged slant angles, and reduces the amount of the pattern that has to be exposed to obtain a tridimensional response. These experiments are performed using patterns of parallel lines.

Cutting and Millard [1984] have performed a quantitative study of the relative importance of size, compression and density gradients in slant judgments of flat as well as curved surfaces. They use textures consisting of disks. By experimenting with conflicting and consistent combinations of different texture

gradients, Cutting and Millard conclude that size and density gradients explain 65% and 28% of the slant judgements of flat surfaces, whereas the compression gradient (gradient of minor-axis length) has practically no effect on the perceived slant. For curved surfaces on the other hand, the compression gradient accounts for almost 96% of the slant judgment with perspective and density gradients having little (8%) impact. The dominance of these selected factors is observed despite the presence of equally strong gradients of other texture features. Thus, it appears that the compression gradient is not important for the perception of a flat surface, but that it is crucial for the perception of curvature. Observers appear to use changes in the compression gradient as a salient local source of information about curvature.

2.5. Distance and foreshortening effects

Two separate effects combine to form the texture gradients observed in an image. Firstly, an increase in the distance between the textured surface and the image plane causes a uniform compression of increasingly large areas of physical texture onto a fixed area of image. Secondly, an increase in foreshortening (the angle between the line of sight and the textured surface) causes an anisotropic compression of the texture. We now turn to a general discussion of the difference in gradients resulting from changing distance and changing foreshortening.

2.5.1. Isotropic effect of changing distance

Texture attributes such as texel shape and texel density undergo an isotropic distortion as the distance between the viewer and the physical texture changes. As the distance to a physical texel increases, the texel subtends a smaller visual angle; a more distant physical texel gives rise to a smaller image texel. This influence of distance on perceived texel extent is isotropic: all dimensions of the texel are scaled equally as distance changes. Therefore, the aspect ratio and the internal angles of the texel are unchanged. Consider an unforeshortened view of a square, for example. The side-length of the square, as measured in the image, depends on the viewing distance; however, the apparent shape of the square -- four sides of equal length, meeting at right angles -- is not affected by the viewing distance.

2.5.2. Anisotropic effect of foreshortening

The effect of foreshortening on apparent texel shape is anisotropic: some dimensions of the texture element shrink more than others. Consider a flat texture where the texels lie in the plane of textured surface (rather than projecting out like porcupine quills). For such a texture, foreshortening is a compression of the texture in the tilt direction. The amount of compression is proportional to $1/\cos(\phi)$, where ϕ is the angle between the line of sight and the textured surface. Foreshortening alters the aspect ratio and internal angles of a texel. For example, a square can foreshorten so that its sides no longer meet at right angles.

2.5.3. Difficulties in interpreting apparent texture density

Apparent texture density is a function of both the distance to the textured surface and the orientation of the textured surface. The effect of distance on texture density is isotropic. However, density has complex behavior under foreshortening: depending on how big the gaps between texels are, the effect of foreshortening may be either isotropic or anisotropic. This is discussed further below.

The simplest characterization of texture density counts the number of texels per unit image-area. In order to measure isotropic versus anisotropic density changes, we may take a set of directional density measurements. We measure the linear density of texels (number of texels crossed per unit distance) along lines at various orientations away from the point of measurement. For texels that fill the plane, as in a brick wall, linear density is easy to compute once texels have been identified. However, we must also

define linear density for sparse textures such as dot patterns (or widely scattered leaves, for example). For sparse textures, linear density may be measured by counting the number of Voronoi polygons crossed per unit distance. A Voronoi polygon associates each point on the plane with the texel that it is closest to.

In a perspective projection, increasing distance shrinks all dimensions equally. If density is measured as the number of texels per unit area, the apparent texture density is proportional to the square of the distance between the camera and the textured surface. The effect of distance on apparent texture density is isotropic. Consider a frontal view of a brick wall for example. (In a frontal view the textured plane is parallel to the image plane.) Draw a horizontal and a vertical line on the image, and count the number of bricks per unit length on each of these lines. Due to the rectangular shape of the bricks, the vertical line has a higher density of bricks than the horizontal line. Suppose that the vertical density is four times as large as the horizontal density. This four-to-one density relationship will be apparent in any frontal view of this brick wall, no matter what the distance to the wall is. Doubling the distance to the wall doubles both the apparent horizontal and vertical texel densities. This unchanging ratio between horizontal and vertical densities illustrates the isotropic effect of distance on texture density.

Since foreshortening causes an anisotropic compression of individual image texels (Section 2.5.2.), it seems intuitively clear that foreshortening must simultaneously have an anisotropic effect on the apparent texture density. This is indeed true for textures composed of plane-filling texels, such as a brick-wall. However, this intuition is false for sparse textures, where the gaps between texels are large relative to the diameters of the texels themselves.

First, consider the foreshortening of plane-filling textures, which behave in an intuitive manner. Consider again a brick wall where each brick is four times as wide as it is high. In a frontal view of this brick wall, the vertical texel density is four times the horizontal texel density. If we foreshorten the wall by rotating it sixty degrees around a vertical axis, we obtain a two-to-one density ratio. On the other hand, if we rotate the wall sixty degrees around a horizontal axis, we obtain an eight-to-one density ratio. Similarly, in an image of a tree trunk the density of texture elements in the vertical direction remains the same in all parts of the trunk, whereas the apparent density in the horizontal direction increases near the edges of the image as the bark curves away from view. For textures such as these, it would be theoretically possible to compare directional densities at two different image points, and decompose the differences into an isotropic scaling component and an anisotropic foreshortening component.

This anisotropic effect of foreshortening on texture density occurs only if the texels are placed adjacent to each other, so that neighbor relations among the texels are preserved during the foreshortening process. In contrast, a texture with small, widely spaced texels experiences a nearly isotropic change in density with foreshortening (although each individual texel is shrunk anisotropically). The sparsest texture possible is a dot pattern, where each texel is a point that occupies no area. Consider a random dot pattern: a subset of points on the plane generated by a Poisson process. An important characteristic of the Poisson process is that the expected number of dots in any region depends only on the area of the region, not on the shape of the region. A slanted view of a random dot pattern of density D results in a random dot pattern of density $D/\cos\theta$, where θ is the angle between the line of sight and the plane of the dot pattern. The slanted dot pattern has an isotropic distribution of dots: a long, thin region can be oriented at any angle without affecting the expected number of points it contains. Thus, an orthographic projection of a slanted dot pattern is not very informative: it is impossible to tell in which direction the dot pattern recedes away from the viewer.

These considerations show some of the difficulties involved in analyzing directional densities. Analysis of changes in apparent texel size and shape seems a more promising approach than analysis of

changes in apparent texel density. Once a shape-based analysis has been performed, texel density measures could be used to verify the results.

2.5.4. Separating distance and foreshortening effects

The appearance of a texture patch is determined by a mixture of perspective and foreshortening effects. Stevens [1981] argues that these two effects need to be separated and discusses methods of doing so (Section 3.9.). In our approach to texture analysis we do not attempt to decompose texture gradients into distance and foreshortening effects. Rather, we hypothesize a particular surface arrangement, compute the total texture distortion (from both distance and foreshortening changes), and then test how well the observed texture gradients in the image match the expectations.

3. PREVIOUS WORK ON INFERENCE OF SURFACE SHAPE FROM TEXTURE

In this section we review some of the computational work on surface estimation from texture. We begin with a summary of the work on representation of homogeneous image texture, and follow this by a review of some approaches to surface estimation from texture.

3.1. Texture representation

A texture has random aspects, and yet appears globally uniform. One of the goals of a texture representation is to characterize the uniformity present in a frontal view of the texture (where the texture sample is parallel to the image plane). In Section 2.3. we discussed the importance of texture uniformity for the recovery of scene layout from non-frontal views of textured surfaces. Texture representations are useful in a variety of other applications, including texture discrimination (for image segmentation), texture recognition, and texture generation (for realism in computer-generated images).

Texture representation is a broad subject which we cannot cover here. We refer the reader to surveys provided by Haralick [1979], Van Gool et al. [1985], and Ahuja and Schachter [1983a], [1983b]. These surveys define two broad classes of texture models: pixel-based statistical models, and region-based structural models. Pixel-based statistical texture measures, such as autocorrelations and cooccurrence probabilities, are useful in texture discrimination and classification applications but do not apply directly to the shape-from-texture problem. Structural texture models focus on the description of texture elements and their placement, and hence are more relevant to the shape-from-texture problem. A texture description that uses independent texel-generation and texel-placement processes provides randomness with overall stable, recognizable characteristics.

3.2. Kender: recovering scene layout from images of man-made textures

Kender [1980a] (alternate references include Kender [1978], [1979], [1983], and Kender and Kanade [1980b]) provides a theoretical framework for shape-from-texture algorithms designed to work with man-made textures. His research was done in the context of analyzing aerial views of cities, where very regular textures, such as sky-scraper windows, provide distance and surface-orientation information. The following topics (among others) are addressed: algorithms for exploiting gravity-based heuristics (the major axes of buildings and trees are aligned with the direction of gravity); and exploitation of texture regularities such as equal-area texels, parallel or perpendicular lines, equal spacing, equal-length lines and symmetry. Kender's main paradigm may be summarized as follows:

- Identify some textural property to "regularize". This property is assumed to be more regular in a frontal view of the texture than in the image. For example, nearly parallel lines in the image may be assumed to originate from precisely parallel lines on the surface.
- Divide the image into significant subimages.
- For each subimage, compute all possible backprojections. Choose the surface orientation that has the most regularized backprojection. (A "backprojection" effectively inverts the foreshortening transformation.)

Kender has efficient methods for precomputing the backprojections for many types of regularization conditions. His method is applicable to regularization conditions relating two texels, such as "nearly equal-

length markings in the image correspond to equal-length markings on the surface". Some texture conditions, such as the one used by Witkin (Section 3.5.), cannot be formulated in this framework. Cues that Kender uses to compute surface orientation include the points of convergence of straight line segments in the image (assuming the physical line segments are parallel; see also Nakatani et al. [1980]), the observed length difference between image line segments (assuming the physical line segments are of equal length) and the observed angle between image line segments (assuming the physical line segments are perpendicular). Kender addresses the issue of perpendicular versus in-plane texture constituents. The precomputed backprojections for perpendicular textures (such as buildings in an aerial view, where each building is a texel) differ greatly from the backprojections appropriate to in-plane texture constituents (such as the windows of a sky-scraper, where each window is a texel). Kender's work, while providing a good framework for the analysis of man-made textures, does not seem applicable to naturally-occurring textures which lack precise regularity in texel spacing, texel size and texel shape.

3.3. Witkin, Davis, Dunn: surface estimation from the observed distribution of edge directions

Witkin [1983] proposes a simple method for estimating surface orientation in orthographic images of natural textures. He assumes that any systematic elongation in a texture is due to foreshortening, and calculates the deprojection that best removes the systematic elongation. The elongation that is present in the image is calculated as follows: (1) apply an edge detector, (2) count the number of edge-elements that occur at each possible edge orientation, (3) calculate which surface orientation would best account for peaks in the edge-orientation histogram (for example, a preponderance of horizontal edge segments suggests that the surface has been rotated around a horizontal axis). Efficient algorithms for performing this calculation are presented by Davis et al. [1983].

Witkin's idea is appealing in its simplicity. However, it is too restrictive to apply to natural images. The image is assumed to be an orthographic projection, so that there are no distortions due to increasing distance from the viewer. Also, the texture must be composed of in-plane texture elements. Witkin's method does not apply well to textures with very non-uniform distributions of edge directions such as checkerboards or herringbone patterns (Davis et al. [1983]). The method fails for elongated textures such as grass, hair, waves, or striated rock: the algorithm attempts to attribute all of the elongatedness to foreshortening, thereby grossly overestimating the slant. Apparently the directional-isotropy assumption is very restrictive and is present only in a small subset of natural images (Aloimonos and Swain [1986, page 585]).

Even when Witkin's assumptions are satisfied, the accuracy of his method is poor. Dunn et al. [1984] describe a series of experiments with implementations of three variations of Witkin's algorithm. The test images are derived from frontal views of textures (from Brodatz [1966]), which are pasted onto cylinders or slanted planes and then digitized. As we point out in Section 4, projections of this type, derived from frontal views, are a simplification of the real projections that results from photographing curved or slanted samples of the physical texture. Even with the simplified projections, slant and tilt estimates obtained from 64-by-64 subwindows are poor. The estimates obtained from 128-by-128 subwindows are better, but large errors still result.

Kanatani [1984] builds on Witkin's work by proposing a different test for the distribution of edge-orientations. He uses an estimator that is based on the number of edge intersections encountered by sets of equi-spaced parallel lines, each set in a fixed direction. If the texture elements have borders with uniform orientation distribution, then the number of intersections is the same for parallel lines in different directions. Otherwise, the observed deviation from a uniform distribution gives an estimate of the surface orientation. This technique is illustrated only on a synthetic example.

3.4. Rosenfeld, Kanatani, Aloimonos: surface estimation from edge density measurements

Rosenfeld [1975] defines a "texture gradient" as the rate and direction of maximum change of texture coarseness across a surface. He suggests measuring the texture gradient by computing the average response of an edge-detection operator in various parts of the image. In coarsely-textured parts of the image there should be fewer edges per unit area than in finely-textured parts of the image. This method assumes that the texture elements do not have significant sub-texture.

Kanatani and Chou [1986] present a theoretical analysis aimed at recovering the 3D shape of a textured surface from a perspective view, assuming that the frontal texture is homogeneous. Dot and line textures are analyzed, to calculate the expected density of dots or lines after perspective projection. The method is illustrated on two synthetic images, one showing a perspective view of a grid of dots and the other showing a perspective view of a grid of lines. The authors do not address the problem of applying this method to real textures.

Aloimonos and Swain [1985], and Aloimonos [1986] describe a procedure to estimate surface shape from measures of texture density; this method has been tested on a wide variety of images. They develop a method that applies when either the number of texels can be counted or the boundaries of the texels can be located. In theory, the orientation of a planar surface can be recovered from the densities measured in two pairs of image regions. Since density fluctuations in the regions can cause inaccurate results, Aloimonos and Swain use a least-square-fit mechanism, which uses density measurements taken from many pairs of image regions. Aloimonos [1986] claims that it is much easier to find the boundaries of texels than to find texels themselves (Section 4 expresses our disagreement with this claim). Therefore he formulates a density measure based on the total length of texel boundaries per unit area. The experimental results reported by Aloimonos are impressive in their scope and accuracy. However, it is our experience that an approach that measures edge density without explicit texel identification cannot work when applied to complex natural textures (with subtexture) under natural lighting conditions (Section 4.1.).

3.5. Ikeuchi: surface estimation from regular patterns

Ikeuchi [1980] proposes a surface estimation algorithm based on the apparent distortion of regular patterns. His method applies only under very restricted conditions. He assumes that the surface texture consists of repetitions of identical texels, and that the frontal shape of the texture element is known. The method is illustrated on synthetic images and on a picture of a golf ball.

3.6. Ohta: computation of vanishing points from observed texel areas

Ohta et al. [1981] propose an interesting method of obtaining the vanishing line of a textured plane from the area of texels in the image. They use the observed areas of pairs of texels to obtain vanishing points. The vanishing points determined by many pairs of texture elements are used to estimate a vanishing line, which gives the direction of tilt. Ohta et al. point out that their method is more general than those described by Kender [1978] and Nakatani et al. [1980] because it does not demand the existence of parallel lines or edges in the texture. However, the method of Ohta et al. has been tested only on synthetic texture images. The problem of extracting texels from natural images is not addressed.

3.7. Zucker: measuring texture coarseness using multi-scale spot detectors

Zucker et al. [1975] suggest a method of measuring texture coarseness. Their goal is to discriminate between a coarse and a fine texture, but a good coarseness-discriminator could also be used to detect texture gradients. Zucker et al. describe a texture discrimination method based on the application of spot

detectors of all different sizes throughout the image. They use a simple spot detector, which computes the difference in average gray-level between two nested, square image-regions. This spot detector does not yield detailed information about spot shapes, but only crude information about spot sizes and spacings. This work is of interest to us, because the goal of the spot detectors is, in effect, to perform texel identification. Our method for texel identification also uses a spot detector (Section 5), one that is more complex and more accurate than Zucker's. Zucker et al. observe that their spot detector is influenced by the presence of subtexture and supertexture.

3.8. Bajcsy and Lieberman: using Fourier transforms to detect texture gradients

Fourier domain features can be used to characterize texture coarseness and elongatedness. Bajcsy and Lieberman [1976] detect texture gradients by calculating Fourier transforms of various parts of the image, determining a characteristic texture-element size from peaks in the Fourier power spectrum, and looking for trends of the characteristic sizes across the image. Their implementation is subject to the following restrictions: (1) texture models are required for choosing appropriate window sizes in which to compute the Fourier transforms (the choice of window size is rather ad hoc, and was manually verified in their experiments), (2) only elongated textures such as grass and ocean waves can be analyzed, and (3) the viewpoint and surface tilt must be known (the texture is assumed to be uniform in a horizontal scan and increasing in density in a bottom-to-top scan, as in an image of a level field of grass). Some of these restrictions are artifacts of the implementation, but there are also difficulties inherent in the use of Fourier spectrum measurements for texture analysis. Natural textures have very irregularly placed texture elements; even in an idealized texture composed of equal-size texture elements, the irregular placement introduces noise into the Fourier spectrum, which obscures the presence of a texture gradient. Also, as discussed by Dyer and Rosenfeld [1976], the choice of window sizes is a very difficult problem in any Fourier-based approach to texture analysis.

3.9. Stevens: separating distance and foreshortening effects

The appearance of a texture patch is determined by a mixture of distance and foreshortening effects (Section 2.5.). Stevens [1981] discusses methods of separating these two effects. He proposes to identify the non-foreshortened dimension of each texel (eg, the major axis of each ellipse in Figure 1). This length depends only on the distance to the texel, and is independent of slant. To find the direction in which to measure these texel widths, the direction of least texture variability must be identified. Successful identification of the non-foreshortened texel dimensions provides the tilt direction as well as the relative distance to each texel.

Surface slant may be obtained either indirectly by differentiation of the estimated distance values, or it may be computed directly from the image. The aspect ratio of the texture elements is a measure which varies with slant and is independent of distance, but Stevens cautions that the relationship between aspect ratio and surface slant is complex. Texels that lie flat on the plane (such as bricks) foreshorten differently than texels that project out of the textured surface (such as erect porcupine quills). Successive occlusion, which occurs for example when one ocean wave partially obscures the view to the next wave, complicates the relationship between aspect ratio and slant even further.

Stevens [1981] presents a good theoretical discussion of the problems involved in defining appropriate texture measures for the extraction of distance and/or surface-orientation information. However, he offers only rather sketchy suggestions for implementation: characteristic dimensions could be estimated from peaks in the Fourier power spectrum, or from measurements of the average distance between edges

provided by $\nabla^2 G$ zero crossings. It appears that these methods will not work on natural textures without extensive modifications.

4. INTEGRATION OF TEXEL IDENTIFICATION AND SURFACE SHAPE ESTIMATION

Texture properties vary across the image in a manner predictable from the physical surface shape; thus it is possible to infer surface shape from texture gradients. This section examines the basic requirements of such an inference process. We argue that correct interpretation of texture gradients requires explicit identification of image texels, especially when texels exhibit significant subtexture. Texture elements cannot be identified in isolation since texels are defined only by the repetitive nature of the texture as a whole. Therefore, we claim, the identification of texture elements is best done in parallel with the estimation of the shape of the textured surface.

4.1. The central role of texture elements

Texture properties are most directly defined in terms of texture elements. Texel identification permits correct analysis of multi-level textures, where texture elements exhibit subtexture. Explicit texel identification also permits a unified treatment of the various texture gradients that may be present in an image. Previous work has avoided texel identification because it is quite difficult to do in real images. Instead, indirect methods are used to estimate texel features. We give below two examples of such methods.

As a first example, consider the edge-based texture features. Edges are normally detected by an edge operator that does not distinguish between texture and subtexture edges, or between edges from different texture fields. Edge density is approximately constant in a frontal view of almost any texture (Aloimonos [1986]); subtexture is not a problem in a frontal view since the same amount of subtexture is visible everywhere. However, a difficulty arises when the texture is seen under projection: more subtexture edges are visible in nearby than in distant samples of the texture. It is incorrect to interpret all the edges produced by an edge detector as the boundaries of texture elements.

This problem is illustrated by Figure 3, which shows all of the edges extracted from several texture images. We use an edge operator described by Nevatia and Babu [1980]. Six 5-by-5 edge masks at different orientations are used; the mask giving the highest output at each pixel is recorded. The edges are thinned through perpendicular non-maximum suppression. The exact details of the edge operator are not important here. We merely wish to illustrate that it would be incorrect to interpret all of the detected edges as boundaries of texture-elements. Additional edges arise due to sub-texture and due to the presence of several texture fields in a single image. The additional edges are not artifacts of this particular edge detector, since they are clearly present in the original images. Many natural textures have a hierarchical physical structure that causes observed edge density to be nearly constant throughout the image: edges from subtexture and sub-subtexture are observed to whatever detail the camera resolution permits.

In the early stages of this research we experimented with measurements of edge density to detect texture gradients. To eliminate sub-texture edges, we experimented with a suppression of weak edges that are located close to strong edges. This is somewhat successful, since the contrast of subtexture is usually less than the contrast of the texture elements themselves. Such edge suppression is an indirect attempt to identify texture elements: the goal is to suppress all edges except those that result from the boundaries of texture elements. We abandoned this edge-based approach in favor of a region-based approach, in which the problem of texel identification is approached more directly, and can thus be solved in a more general way.

As a second example of indirect estimation of texture features, consider the methods that make specific assumptions about the distribution of texel edge directions. Witkin [1981], Dunn [1984] and Kanatani [1984] assume that the texel edges in a frontal view of the texture have an isotropic directional distribution, which may not be true for many textures. Kender [1979], [1980a], [1983], Kender and Kanade [1980b], and Nakatani et al. [1980] consider textures containing parallel or perpendicular lines; these include many man-made textures but few natural textures. All of the edge-direction methods, like the edge-density methods, are sensitive not only to edges arising from texel borders, but also to edges arising from subtexture and from multiple texture fields.

Texture algorithms are often tested on images formed by artificial projections derived from images of frontal texture samples. An artificial projection is formed in one of two ways. The first method is to wrap an image of the frontal texture onto a surface such as a slanted plane or a cylinder; a view of this surface is then digitized to obtain the test image (see, for example, Dunn et al. [1984]). The second method obtains a similar result using a computer program. Starting with a digitized sample of a real texture seen in frontal view, the computer program applies a perspective transformation to map the digitized texture sample onto a desired surface geometry. Both of these methods produce simplified approximations of the images that result when curved or slanted samples of physical texture are photographed and digitized. Artificial projections lose the effect of three-dimensional relief: texels do not shadow or occlude each other, and they may foreshorten improperly (imagine the result of performing a synthetic projection of erect porcupine quills). Most importantly, artificial projections do not properly capture the complexities of subtexture: no subtexture details appear when regions of the frontal texture sample are expanded to model parts of the surface that are close to the viewer. Since artificial projections introduce these simplifications, texture algorithms that successfully analyze artificially-projected scenes cannot necessarily cope with real images of slanted physical textures.

We summarize with the following observations. By making some assumptions about the nature of texture elements it is often possible to estimate certain texel properties through measures that do not require explicit identification of texture elements. However, when texture elements are not identified and explicitly dealt with, it becomes difficult to distinguish between responses due to texture elements and those due to other image features. Edge-density measurements (Section 3.6.) may include contributions from subtexture or supertexture edges, from borders of partially occluded texture elements, and from edges of texels belonging to several texture fields. Similarly, when making an edge-direction histogram (Section 3.5.) it may not be possible to distinguish between edges from texel borders and edges due to other features such as subtexture. Fourier domain features (Section 3.8.) are also sensitive to the presence of subtexture and supertexture. It appears to be necessary to recognize the texture elements before the various measures can be computed as intended.

Explicit identification of texture elements offers an additional advantage: texture elements provide a unifying framework for examination of the various texture gradients that may be present in an image. The relative accuracy of texture gradients varies from image to image (Section 2.3.); therefore it is not known in advance which texture gradients can be measured accurately enough to be useful for the estimation of three-dimensional scene layout. A long-term goal of our research is to provide a unified treatment of various texture gradients. The current implementation, summarized in Section 8, is only a start in this direction: we use the area-gradient of texture elements and the area-gradient of the spaces between the texture elements.

4.2. Integration of texel identification and surface shape estimation

Having said that we must identify image texels, we now address the problem of texel identification. We claim that texel identification is best done in parallel with the estimation of surface shape. To see this, consider the image regions of relatively uniform gray level. Regions of relative gray-level uniformity arise in many different ways. An image region may correspond to a texture element, or to the visible portion of a partially occluded physical texel. Alternately, the region may represent subtexture within a close-range texture element or supertexture arising from a merging of several texture elements located at large distances in the scene. Finally, the region may arise from an isolated object that is not part of a texture (eg, the snowy areas and the tree trunk in the rock-pile image of Figure 5).

If we consider a single image region in isolation, it is impossible to tell to which texture field, if any, the region belongs. This decision can only be made by considering the rest of the image: could this region be a texel that is consistent with the properties of many other image texels? To answer this question we must hypothesize a surface estimate. It is therefore essential that the identification of texture elements and the estimation of surface shape be done cooperatively.

We have developed a two-step approach to carry out such integration of texel identification and surface estimation. First, we assume that all homogeneous gray-level regions are candidates for being texels; thus the first step performs a local gray-level analysis to identify potential texels. Second, we use surface-fitting to identify the true texels from among the candidates, while simultaneously constructing an approximation to the shape of the textured surface. The second step thus enforces perspective viewing constraints to select texels. The next three sections describe the algorithm that we have implemented. Section 5 describes a region detector for extracting uniform image regions of unknown size and shape. Derivations necessary for the surface-fitting are presented in Section 6, and the surface-fitting algorithm is described in Section 7. Section 8 contains a summary of the implementation, and presents results for a variety of images of textured natural scenes.

5. MULTISCALE EXTRACTION OF HOMOGENEOUS IMAGE REGIONS

It is nearly impossible to extract texture elements directly from an image because of the tremendous variety among textures and because the apparent size and shape of texture elements varies across the image. We decompose this problem into two more tractable parts: first we extract a large set of candidate texels from the image, and then we select among these candidates to find a set of texels that shows variations consistent with a particular three-dimensional surface arrangement. This section describes the extraction of candidate texels; later sections describe a method for selecting texels from among the candidates, while simultaneously finding the three-dimensional surface arrangement.

Any region that has relatively uniform gray-level is a candidate texel. The uniformity of small regions is measured relative to a small surrounding neighborhood in the image, whereas the uniformity of large regions is measured relative to a proportionally larger neighborhood in the image. Since the shape and size of texture elements is unknown in advance, we need a multiscale operator for detecting uniform regions of all shapes and sizes. We simplify this problem by assuming that each region can be represented as a union of overlapping circular disks. Large disks define the rough shape of the region, with overlapping smaller disks capturing finer shape details such as protrusions and concavities. We present a multiscale method of extracting all circular image regions of relatively uniform gray level. Sets and subsets of overlapping disks are used to form candidate texture elements.

5.1. Scale space

The region-extraction algorithm is based on an analysis of the scale-space behavior of uniform image regions. Before presenting a derivation of the algorithm we briefly review previous research concerning multi-scale image representations.

The term *scale space* was introduced by Witkin [1983]. He builds on the theory of edge detection developed by Marr and Hildreth [1980] (see also Marr [1982]), in which edges are located as the zero-crossings in the Laplacian of a Gaussian-smoothed image. Marr and Hildreth suggest using a selection of filter sizes in order to capture edges at different scales: thin, sharp edges are best captured by small filter sizes whereas broad, fuzzy edges are better characterized by large filter sizes. However, Marr and Hildreth do not adequately address the problem of combining the edge images obtained from various filter sizes. Witkin [1983] introduces a scale-space representation of $\nabla^2 G$ zero-crossings over a continuous range of scales. A scale-space representation is constructed by convolving the original signal with $\nabla^2 G$ filters for all possible choices of the filter size σ. The scale-space representation of a one-dimensional signal occupies an x–σ plane, whereas the scale-space representation of a two-dimensional signal (such as an image) occupies an x–y–σ volume. Gaussian smoothing has two effects: simplification through removal of fine-scale features, and distortion through dislocation, broadening and flattening of the surviving features. Salient zero-crossing contours may be identified at coarse scales, and then traced to fine scales for accurate localization. Witkin [1983] describes an efficient representation of the zero-crossings of a one-dimensional signal (in the $x-\sigma$ plane). There is no straightforward extension of this representation to encode the zero-crossings of two-dimensional signals.

Crowley and Parker [1984] analyze images over a range of scales using a representation that is related to Witkin's scale-space representation. Crowley and Parker use a difference-of-Gaussian operator, which may be considered a discrete approximation to the $\nabla^2 G$ operator. (The relationship between $\nabla^2 G$

and the difference of two Gaussians is characterized by the diffusion equation $\nabla^2 G = \frac{1}{\sigma}\frac{\partial}{\partial\sigma}G$. The difference of two Gaussians with similar σ is a discrete approximation to $\frac{\partial}{\partial\sigma}G$ and hence to $\nabla^2 G$.) A scale-space representation of a signal has many features that could be analyzed. Whereas Witkin concentrates on the behavior of zero-crossings over a range of scales, Crowley and Parker instead concentrate on peaks and ridges extracted over a range of scales. A peak in the $\nabla^2 G$ response indicates a local best-fit of a disk of a particular size. The pattern of peaks and their connecting ridges characterizes object shapes in a form that is suited to object recognition or matching: the coarse shape information captured by the large filter sizes is used to bring the objects into approximate registration, and then the more detailed shape information captured by the small filter sizes is used to refine the matching.

We have developed a method of analyzing the scale-space behavior of an image to extract primitive shapes that together span the image regions. A description of our method follows.

5.2. Notation

The following symbols are used:

∇	*gradient:* $\nabla F = (\frac{\partial F}{\partial x}, \frac{\partial F}{\partial y})$
∇^2	*laplacian:* $\nabla^2 F = \frac{\partial^2 F}{\partial x^2} + \frac{\partial^2 F}{\partial y^2}$
G	*unnormalized Gaussian:* $e^{-r^2/2\sigma^2}$ (where $r = \sqrt{x^2+y^2}$)
G_n	*normalized Gaussian:* $\frac{1}{2\pi\sigma^2} e^{-r^2/2\sigma^2}$
$\nabla^2 G$	*laplacian of unnormalized Gaussian, positive center lobe:* $\frac{2\sigma^2 - r^2}{\sigma^4} e^{-r^2/2\sigma^2}$
$\nabla^2 G_n$	*laplacian of normalized Gaussian, positive center lobe:* $\frac{2\sigma^2 - r^2}{2\pi\sigma^6} e^{-r^2/2\sigma^2}$
C	*contrast of a bar or disk*
B	*bar width*
D	*disk diameter*
σ, w	*G or* $\nabla^2 G$ *filter size;* $w = 2\sqrt{2}\sigma$

The various forms of the $\nabla^2 G$ operator used in the literature differ from each other by a multiplicative constant. Multiplicative constants do not change the shape of the $\nabla^2 G$ operator (Grimson and Hildreth [1985]); however, since they do alter the shape of the $\frac{\partial}{\partial\sigma}\nabla^2 G$ operator, we make the distinction between $\nabla^2 G$ and $\nabla^2 G_n$. Differences between the $\nabla^2 G$ and $\nabla^2 G_n$ operators are discussed further in Section 5.4.

In keeping with tradition in the literature, we negate the $\nabla^2 G$ equations, so that filters with a positive center lobe result. The size of a $\nabla^2 G$ filter is characterized by σ, the standard deviation of the Gaussian distribution, or by w, the width of the positive center lobe of the $\nabla^2 G$ filter.

5.3. Closed form expressions for the ∇^2G responses of disk and bar images

Our algorithm for uniform-region extraction is based on calculations of the $\nabla^2 G$ and $\frac{\partial}{\partial\sigma}\nabla^2 G$ responses of a disk image.

Definition: Given a function $I(x, y)$ which specifies the intensity of an image, the **∇^2G response** of this image at (x, y) is given by the following convolution:

$$\nabla^2 G(x,y) \;*\; I(x,y) = \int_{-\infty}^{+\infty}\int \frac{2\sigma^2-(u^2+v^2)}{\sigma^4}\; e^{-(u^2+v^2)/2\sigma^2}\; I(x-u, y-v)\, du\, dv \qquad (5.1)$$

This definition is for continuous rather than discrete images. We analyze the $\nabla^2 G$ response of ideal disks and bars in this continuous domain. (However, to generate the $\nabla^2 G$ convolution of digitized images, we sample the $\nabla^2 G$ filter values and perform a discrete convolution.) Mathematical analysis of the response of the $\nabla^2 G$ filter to most images is difficult because the convolution integrals of Equation (5.1) do not have closed form solutions. However, a closed-form solution can be derived for the center point of a circular disk of constant intensity. The image of a disk of diameter D and contrast C is defined by

$$\text{disk image:} \qquad I(x,y) = \begin{cases} C & \text{if } x^2+y^2 \le D^2/4 \\ 0 & \textit{elsewhere} \end{cases} \qquad (5.2)$$

Using this definition of $I(x,y)$ in Equation (5.1), and setting x and y to zero, we show in the appendix that the $\nabla^2 G$ response at the center of the disk is $\frac{\pi C D^2}{2\sigma^2} e^{-D^2/8\sigma^2}$ and the $\frac{\partial}{\partial\sigma}\nabla^2 G$ response at the center of the disk is $\frac{\pi C D^2}{2}\left(\frac{D^2}{4\sigma^5} - \frac{2}{\sigma^3}\right) e^{-D^2/8\sigma^2}$.

We also solve Equation (5.1) for an image of an infinitely-long bar. An infinitely long bar is not a useful shape primitive; however, the bar response is used for calculations performed in Sections 5.4. and 5.9. The image of a bar of width B and contrast C is defined by

$$\text{bar image:} \qquad I(x,y) = \begin{cases} C & \text{if } 0 \le x \le B \\ 0 & \textit{elsewhere} \end{cases} \qquad (5.3)$$

Using this definition of $I(x,y)$ in Equation (5.1), and setting x to $B/2$, we show in the appendix that the $\nabla^2 G$ response at the center of the bar is $\frac{\sqrt{2\pi} CB}{\sigma} e^{-B^2/8\sigma^2}$ and the $\frac{\partial}{\partial\sigma}\nabla^2 G$ response at the center of the bar is $\sqrt{2\pi} CB\left(\frac{B^2}{4\sigma^4} - \frac{1}{\sigma^2}\right) e^{-B^2/8\sigma^2}$.

TABLE 1

	center of bar	center of disk
$\nabla^2 G$	$\frac{\sqrt{2\pi} CB}{\sigma} e^{-B^2/8\sigma^2}$	$\frac{\pi C D^2}{2\sigma^2} e^{-D^2/8\sigma^2}$
$\nabla^2 G_n$	$\frac{CB}{\sqrt{2\pi}\sigma^3} e^{-B^2/8\sigma^2}$	$\frac{CD^2}{4\sigma^4} e^{-D^2/8\sigma^2}$
$\frac{\partial}{\partial\sigma}\nabla^2 G$	$\sqrt{2\pi} CB\left\{\frac{B^2}{4\sigma^4} - \frac{1}{\sigma^2}\right\} e^{-B^2/8\sigma^2}$	$\frac{\pi C D^2}{2}\left\{\frac{D^2}{4\sigma^5} - \frac{2}{\sigma^3}\right\} e^{-D^2/8\sigma^2}$
$\frac{\partial}{\partial\sigma}\nabla^2 G_n$	$\frac{CB}{\sqrt{2\pi}}\left\{\frac{B^2}{4\sigma^6} - \frac{3}{\sigma^4}\right\} e^{-B^2/8\sigma^2}$	$\frac{CD^2}{4}\left\{\frac{D^2}{4\sigma^7} - \frac{4}{\sigma^5}\right\} e^{-D^2/8\sigma^2}$

Table 1 summarizes the expressions derived in the appendix. The correctness of these equations has been verified experimentally by performing discrete convolutions of $\nabla^2 G$ and $\frac{\partial}{\partial\sigma}\nabla^2 G$ masks with synthesized images of isolated bars and disks. The $\nabla^2 G$ and $\frac{\partial}{\partial\sigma}\nabla^2 G$ values at the centers of the bars and disks match the values predicted by the equations to within roundoff and discretization errors.

5.4. ∇^2G convolutions have more consistent magnitude than $\nabla^2 G_n$ convolutions

If we are considering the $\nabla^2 G$ responses of a disk or bar, there are two quantities we may vary: the filter size (σ) and the disk-diameter (D) or bar-width (B). This is illustrated in Figure 2, which shows the $\nabla^2 G$ response to a scaled square wave over a range of σ values. The square wave is a cross section of an image composed of infinitely long bars of varying widths. A horizontal scan of Figure 2 shows the dependence of the $\nabla^2 G$ response on bar width. A vertical scan of Figure 2 shows the response of a fixed-width bar to $\nabla^2 G$ filters of varying size. Observe that with fixed σ there is an ideal bar-width that gives maximal response. Similarly, with fixed bar-width there is an ideal filter-size that gives maximal response. We would like these ideal bar-widths and ideal filter-sizes to coincide. The desired consistency property for the $\nabla^2 G$ magnitudes seen at a bar center is:

> If $\hat{B}$ is the width of the maximally responding bar at a fixed $\hat{\sigma}$, then, conversely, $\hat{\sigma}$ is the filter size that maximizes the response of a bar of width $\hat{B}$.

Similarly, the desired consistency property for the $\nabla^2 G$ magnitudes seen at a disk center is:

> If $\hat{D}$ is the diameter of the maximally responding disk at a fixed $\hat{\sigma}$, then, conversely, $\hat{\sigma}$ is the filter size that maximizes the response of a disk of diameter $\hat{D}$.

Using the equations in Table 1, it is easy to show that these consistency properties hold for $\nabla^2 G$, but not for $\nabla^2 G_n$. In order to prove this, we find the σ values which maximize the response for fixed-size bars and disks, and compare this to the bar and disk sizes which maximize the response at a fixed σ. We set $\frac{\partial}{\partial B}$ of the $\nabla^2 G$ bar-center response and $\frac{\partial}{\partial D}$ of the $\nabla^2 G$ disk-center response to zero to find that

> for fixed σ, both $\nabla^2 G$ and $\nabla^2 G_n$ have maximum response to disks of diameter $D = 2\sqrt{2}\sigma$ and to bars of width $B = 2\sigma$.

By setting the $\frac{\partial}{\partial \sigma}\nabla^2 G$ and $\frac{\partial}{\partial \sigma}\nabla^2 G_n$ expressions from Table 1 to zero, we find that

> for a fixed disk-diameter D (varying σ), $\nabla^2 G$ has maximum response at the disk center when the filter size $\sigma = D/(2\sqrt{2})$; for a fixed bar-width B, $\nabla^2 G$ has maximum response at the bar center when the filter size $\sigma = B/2$.

On the other hand,

> for a fixed disk-diameter D (varying σ), $\nabla^2 G_n$ has maximum response at the disk center when the filter size $\sigma = D/4$; for a fixed bar-width B, $\nabla^2 G_n$ has maximum response at the bar center when the filter size $\sigma = B/(2\sqrt{3})$.

The consistency property for $\nabla^2 G$ follows by inspection; this property is useful for comparing the responses of an image to $\nabla^2 G$ filters of various sizes.

5.5. Estimating the size and contrast of disks and bars from ∇^2G measurements

We have seen that $\nabla^2 G$ responds maximally to disks of diameter $D = 2\sqrt{2}\sigma$. Imagine an image composed of non-overlapping equally-bright disks of many different sizes. The $\nabla^2 G$ response at some particular σ will peak for particular disk sizes, namely for those disks with diameters close to $2\sqrt{2}\sigma$. This effect is illustrated for a one-dimensional signal by Figure 2, which shows the $\nabla^2 G$ response to a scaled square wave over a range of σ values. (Since the square wave is a cross section of an image composed of infinitely long bars, the $\nabla^2 G$ response in Figure 2 peaks for those square pulses with widths close to 2σ.) It seems possible to characterize image structure by noting the values of local maxima in the $\nabla^2 G$ response at various values of σ. For small σ, small regions give maximum response; for larger σ, larger regions give maximum response. However, we need additional measurements in order to distinguish high-contrast regions from large regions. The $\nabla^2 G$ response at a disk center depends on both the disk-diameter D and

the disk-contrast C (Table 1); therefore, many different disks can give the same $\nabla^2 G$ response at a given σ. To avoid interpreting lighting changes as region-size changes, the image response to $\nabla^2 G$ filters of various sizes must be analyzed.

The equations in Table 1 suggest how to compensate for the influence of region contrast. Both $\nabla^2 G$ and $\frac{\partial}{\partial\sigma}\nabla^2 G$ responses are proportional to the region contrast C; dividing one value by the other leads to a measure independent of C. From the entries in Table 1, we see that at the center of an ideal circular disk,

$$(\tfrac{\partial}{\partial\sigma}\nabla^2 G*I)/(\nabla^2 G*I) = \frac{D^2}{4\sigma^3} - \frac{2}{\sigma}$$

We can solve this equation for the disk diameter D:

$$D = 2\sigma\sqrt{\sigma(\tfrac{\partial}{\partial\sigma}\nabla^2 G*I)/(\nabla^2 G*I) + 2} \tag{5.4}$$

where the convolutions are evaluated at the center of the disk. Once we have solved for the disk-diameter, we obtain the contrast C by

$$C = \frac{2\sigma^2}{\pi D^2}\, e^{D^2/8\sigma^2}\, (\nabla^2 G*I) \tag{5.5}$$

where the convolution is evaluated at the center of the disk. Similarly the bar-width B may be calculated as

$$B = 2\sigma\sqrt{\sigma(\tfrac{\partial}{\partial\sigma}\nabla^2 G*I)/(\nabla^2 G*I) + 1} \tag{5.6}$$

where the convolutions are evaluated at the center of the bar. The $\nabla^2 G$ response for a particular disk or bar is maximized when $\frac{\partial}{\partial\sigma}\nabla^2 G*I = 0$; at this point Equations (5.4) and (5.6) evaluate to $D = 2\sqrt{2}\sigma$ and $B = 2\sigma$.

5.6. Detecting uniform regions in real images

In the previous section we derived the theoretical results necessary for the definition of a region-extraction algorithm. The algorithm is based on Equations (5.4) and (5.5): after computing the discrete convolution of a real image with $\nabla^2 G$ and $\frac{\partial}{\partial\sigma}\nabla^2 G$ masks, Equations (5.4) and (5.5) are applied at selected image locations to recover the diameters and contrasts of the disks that best fit the local shape of uniform image regions. The equations must be applied at disk centers; the equations produce nonsensical results at image locations where the intensity pattern is not at all disk-like. (If the quantity under the the square root symbol in Equation (5.4) is negative an attempt was made to apply the equations at an unsuitable location; no disk can be fit.) Since the suitable locations for disk centers are not known a priori, Equations (5.4) and (5.5) are applied at all pixels that are local maxima in the $\nabla^2 G$ image. This produces disks that model the positive-contrast regions in the image; the equations are also applied at local minima to obtain disks that model the negative-contrast regions in the image. The generality of this region detector is discussed in Section 5.9.

Choosing local maxima of $\nabla^2 G$ as potential disk centers is justified by the following considerations. Consider a near-circular image region of approximately uniform gray-level. Local maxima in the $\nabla^2 G$ image occur at the region center for any filter size that is close to the diameter of the region. However, as illustrated in Figure 2, if σ is chosen much too small or much too large, then the $\nabla^2 G$ local maxima do not locate the region center well. If σ is too small, then the local maxima occur off-center (and application of Equation (5.4) underestimates the region diameter). On the other hand, if σ is too large, then Gaussian smoothing merges neighboring regions, making the result of Equation (5.4) meaningless.

Thus a selection of filter sizes is necessary to assure that at least one of the filter sizes falls into the σ range at which it is appropriate to analyze the local shape of each region. We apply Equations (5.4) and (5.5) at $\nabla^2 G$ local maxima for six different σ values. A disk detected at a filter size σ is accepted only if $2\sqrt{2}\sigma$ (the diameter of the center lobe of the $\nabla^2 G$ filter) is close to the disk diameter. Other disks are located more accurately at another filter size. Implementation details are covered in Section 5.8.1.

5.7. Forming candidate texture elements from groups of overlapping disks

Homogeneous regions in an arbitrary texture have complex shapes. We construct an approximation of these complex shapes using a union of overlapping circular disks. After all disks have been detected for a particular image, overlapping disks are used to form potential texture elements. When overlapping disks are grouped together, concavities are formed at the joins between the disks. At each concavity, we can choose either to keep the complete set of disks, or to split into two smaller sets of disks. The significance of a concavity is not always clear. Some concavities arise at the border between two neighboring texels; at other times the concavities are part of the shape of an individual texel. Since there is no a priori way to tell which set of disks (split or unsplit) is a better representation of a texel, all possible sets of disks are added to the list of candidate texels. When a disk participates in the formation of several candidate texels, these candidates are marked as mutually exclusive, so that at most one of them is accepted as a true texel. Details of the implementation are covered in Section 5.8.2.

5.8. Implementation details for the region detector

Let I denote an image. The processing of I is divided into three main phases: finding the disks, constructing potential texture elements from the disks, and fitting a planar surface to the candidate texels. Here we discuss the implementation of the first two phases. Implementation of the third phase is described in Section 7. Figures 5 to 38 show the positive-contrast and negative-contrast regions extracted from various images. Figure 4 illustrates details of the disk-fitting process for one particular image.

5.8.1. Finding disks

The first step in processing an image I is to compute $\nabla^2 G * I$ and $\frac{\partial}{\partial\sigma}\nabla^2 G * I$ for a selection of filter sizes. To compute $\nabla^2 G * I$ for a particular σ value, the image is convolved with a mask whose coefficients are taken from

$$\frac{2\sigma^2 - r^2}{\sigma^4} e^{-r^2/2\sigma^2}$$

To compute $\frac{\partial}{\partial\sigma}\nabla^2 G * I$ for a particular σ value, the image is convolved with a mask whose coefficients are taken from

$$\frac{6r^2\sigma^2 - r^4 - 4\sigma^4}{\sigma^7} e^{-r^2/2\sigma^2}$$

The convolutions are performed via multiplication in the Fourier domain. Six different $\nabla^2 G$ and $\frac{\partial}{\partial\sigma}\nabla^2 G$ convolutions are evaluated, using σ values of $\sqrt{2}$, $2\sqrt{2}$, $3\sqrt{2}$, $4\sqrt{2}$, $5\sqrt{2}$ and $6\sqrt{2}$. The center lobes of the six $\nabla^2 G$ filters have diameters of 4, 8, 12, 16, 20 and 24 pixels respectively.

The second step in processing the image I is to mark the locations where disks will be fit. In order to find disks that model positive-contrast image regions, each $\nabla^2 G * I$ image is scanned to find local maxima: any pixel larger than all eight of its neighbors is marked as a disk-center location. Similarly, in order to find disks that model negative-contrast image regions, each $\nabla^2 G * I$ image is scanned to find local

minima: any pixel smaller than all eight of its neighbors is marked as a disk-center location.

Next, Equations (5.4) and (5.5) are applied at each of the marked locations, using the $\nabla^2 G^* I$ and $\frac{\partial}{\partial\sigma}\nabla^2 G^* I$ values observed at that location:

$$D = 2\sigma\sqrt{\sigma(\tfrac{\partial}{\partial\sigma}\nabla^2 G^* I)/(\nabla^2 G^* I) + 2} \qquad C = \frac{2\sigma^2}{\pi D^2}\, e^{D^2/8\sigma^2}\,(\nabla^2 G^* I)$$

Disks are detected most accurately at a filter size close to their diameter ($D \simeq w = 2\sqrt{2}\sigma$); therefore only a restricted range of disks diameters is accepted from each filter size. In the current implementation, the detected disk diameter must be within two pixels of the filter size. Thus, of the disks detected by the filter of width 12 pixels ($\sigma = 3\sqrt{2}$), we keep only those with diameters in the range 10 to 14 pixels. Internally, the disks are represented as a list of disk-descriptors, where each disk-descriptor contains the coordinates of the disk center, the disk diameter and the disk contrast. However, for display purposes the disks may be expanded to fill the regions they represent. Parts (d) to (k) of Figure 4 illustrate the disks detected in an image of a rock pile at various filter sizes. In Figure 4 each disk is represented with an intensity proportional to its contrast. Note that the smaller filter sizes find many more disks than the larger filter sizes do: the expected distance between $\nabla^2 G^* I$ zero-crossings is proportional to σ (Marr [1982], page 136), and hence the density of local maxima (or minima) is proportional to $1/\sigma^2$.

The final step in disk detection is to combine the disks detected at the various filter sizes into one data structure. This is done by concatenating the lists of disk-descriptors from each filter size. Parts (b) and (c) of Figure 4 show the result. Only one disk can be displayed at pixel locations covered by several disks. In part (b) the disk of higher contrast is displayed; therefore, low-contrast disks that are spatially contained within high-contrast disks are not visible. Part (e) shows the low-contrast disks better: at pixel locations covered by several disks, the disk of lower contrast is displayed.

5.8.2. Constructing potential texture elements from the disks

After the disks have been detected, overlapping disks are grouped to form a list of potential texture elements. We process one group of overlapping disks after another, extracting all subsets of disks that are spatially connected and contain no concavities greater than 90°. Concavities are computed as the angle formed between two neighboring disks on the border of a region. A concavity greater than 90° forces a split into smaller regions. A concavity in the range 50° to 90° causes both the unsplit and split regions to be included on the list of potential texels. Concavities less than 50° are never split. If a concavity is in the range 50° to 90°, the disks are used to form three potential texture elements: one large region consisting of all the disks, and two smaller regions resulting from splitting the large region at the concavity[1]. These rules are applied recursively, so that the smaller regions can again give rise to several alternate entries on the list of potential texture elements. The particular values 50° and 90° are not critical; we have found that the range 50° to 90° is large enough to capture all regions of interest and yet small enough to prevent a combinatorial explosion in the number of potential texture elements generated. Potential texture elements that share a disk are marked as mutually exclusive, so that at most one of them can contribute support to a planar fit and be chosen as a true texture element.

---

[1]Region splitting is implemented as follows. We begin with a set P of overlapping disks, which together cover an image region R. The largest concavity in R is found by computing the angles formed by every pair of neighboring disks on the border of R. Suppose that X and Y are two neighboring disks on the border of R, and that they form a concavity that should cause a split into smaller, more convex regions. The concavity is split by (1) removing X from P and repeating the above process, and then (2) removing Y from P and repeating the above process.

Internally, each potential texture element is represented by a texel-descriptor. A texel-descriptor contains a list of the disks that together represent the image region occupied by the texel. The texel-descriptor also contains other information, including the area, average gray-level and contrast of the texel, as well as a list of texels that are mutually exclusive with this one.

5.9. Generality of the representation

The region detector described above performs well on a wide variety of images. Parts (a) and (b) of Figures 5 to 38 illustrate the strengths and weaknesses of the region extraction. The most notable weakness of the region extraction is that the representation of elongated regions is not very good. This is not surprising, since the only shape primitive used is a circular disk. In Section 8.2. we mention future work that could lead to the development of additional shape primitives more suited to the detection of elongated regions. Here we analyze the result produced by our region detector when it is applied to an elongated image region. Two sources of error are apparent: (1) the calculated disk diameters overestimate the widths of elongated regions, and (2) long thin texels tend to appear as a string of disconnected disks. We discuss these two types of errors in turn.

Suppose we have an image of an infinitely long bar of width B, and we try to fit a disk to some point along the center-line of the bar. By comparison of Equations (5.4) and (5.6) we can calculate the diameter of the resulting disk. The disk diameter will overestimate the bar-width since the formula

$$D = 2\sigma\sqrt{\sigma(\tfrac{\partial}{\partial\sigma}\nabla^2G^*I)/(\nabla^2G^*I) + 2}$$

is used to obtain the diameter D that models the bar-width B, whereas the correct formula for the bar is

$$B = 2\sigma\sqrt{\sigma(\tfrac{\partial}{\partial\sigma}\nabla^2G^*I)/(\nabla^2G^*I) + 1}$$

The seriousness of this error depends, of course, on the magnitude of $\sigma(\frac{\partial}{\partial\sigma}\nabla^2G^*I)/(\nabla^2G^*I)$ relative to 1. In our implementation, the quantity $(\frac{\partial}{\partial\sigma}\nabla^2G^*I)/(\nabla^2G^*I)$ is small. We accept a disk detected at a particular filter size only if the diameter is close to the filter size: $D \approx w = 2\sqrt{2}\sigma$. When $D = 2\sqrt{2}\sigma$ we have $(\frac{\partial}{\partial\sigma}\nabla^2G^*I)/(\nabla^2G^*I) = 0$, so the calculated disk diameter overestimates the bar width by a factor of $\sqrt{2}$. Thus, in an image of an infinitely long region, the region-width is overestimated by a factor of approximately $\sqrt{2}$. For regions that are more moderately elongated the overestimation is less serious. In the limiting case of a region with no elongation, there is no overestimation at all.

Using a circular disk as a shape primitive, we hope to model elongated regions by a string of overlapping disks. However, in our current implementation the disks that model an elongated region are often placed too sparsely, so that a disconnected chain of disks results. One possible remedy is to fit disks more closely. Currently we fit disks at local maxima (or minima) of the ∇^2G^*I images. An elongated region gives rise to a ridge of large values in the ∇^2G^*I image. Such a region could be better modeled by fitting a disk at each ridge point rather than just at each local maximum. However, it is difficult to formulate simple and accurate criteria for judging when a ridge point is significant enough to be used as a disk center.

Inaccurate modeling of elongated regions does not necessarily cause errors in the analysis of textures composed of elongated texels. In the present implementation we use the detected regions to analyze gradients of texel area. All elongated regions in an image are split into a chain of disks in a statistically similar way; thus we successfully analyze images of elongated textures (see, for example, Figures 29 and 30), even though the image texels we identify are not as elongated as they should be. A better extraction of elongated regions would allow us to detect additional texture gradients based on other texture features. For

example, the aspect ratio of the detected texels could be analyzed. In the current implementation the measured aspect ratios are too inaccurate to be informative.

The region detector is designed to respond to regions of relatively uniform gray level that contrast with a relatively uniform background. Most of the regions in the images of Figures 5 to 38 satisfy these conditions (at least approximately). Therefore region detection is quite good. However, it is not difficult to construct images containing uniform regions to which our region detector does not respond. Consider, for example, an image which is white on the left half and black on the right half, with a gray region centered on the border between black and white. Our detector will not respond well to the gray region because the background around the region is highly nonuniform.

In summary, despite its shortcomings, the region detector is exact enough to allow fairly accurate detection of the gradient of texel area in the images shown in Figures 5 to 38. Our method of detecting and modeling the texture gradients is described in the next two sections.

6. TEXTURE GRADIENTS FOR PLANAR SURFACES

In order to deduce scene-layout from texture cues, we must first quantify the relationship between 3D scene layouts and the corresponding texture gradients. In this section we analyze the texture gradients present in images of planar textured surfaces. The analysis applies to textures that have no three-dimensional relief, such as the texture of a wooden table top. The texture gradients are characterized by deriving the relationship between physical texels and image texels as a function of image location, scene layout, and camera parameters.

The results of this section can be summarized with reference to a perspective view of an idealized disk texture as in part (a) of Figure 1. Scanning this image from left to right, in the direction of constant surface depth, no texture gradient is observed. All image-texels encountered in a horizontal scan have the same size and shape. On the other hand, scanning the image from bottom to top, in the direction of greatest depth increase, changes in the size and shape of image texels are observed. We characterize the magnitude of the observed changes in image texels as follows:

1. The length of the major axes decreases linearly as the image is scanned from bottom to top. This is a distance-scaling effect, due to the changing distance between the physical texels and the viewer.
2. The length of the minor axes decreases quadratically as the image is scanned from bottom to top. This quadratic decrease occurs because the minor axes are subject to foreshortening as well as to distance scaling. (Foreshortening is proportional to the angle between the line of sight and surface normal.)
3. Texel area is proportional to the product of the major and minor axis lengths. Therefore, the texel areas decrease cubically as the image is scanned from bottom to top.

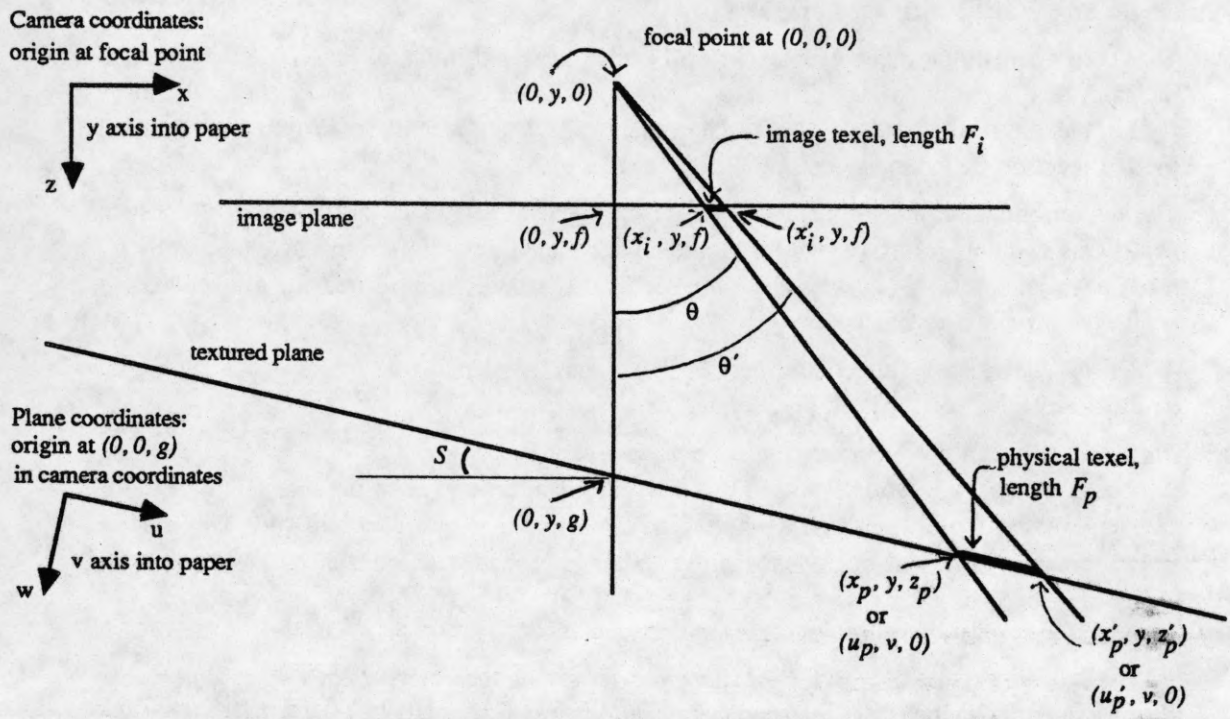
Additional conclusions may be drawn about the rate of change of texel eccentricity and texel density. Since eccentricity equals the ratio of major axis length to minor axis length, the texel eccentricity increases linearly as the image is scanned from bottom to top. The density of texels (number of texels observed per unit area in the image) increases cubically as the image is scanned from bottom to top.

6.1. Notation for scene layout and camera geometry

We consider a planar surface covered with a pattern of identical texels. Later in this section two expressions are derived to describe the size of image texels. The first expression characterizes the texel extent in the direction of greatest depth increase (the minor axis length in Figure 1). The second expression characterizes the texel extent in the direction of no depth change (the major axis length in Figure 1). Combining these we derive an expression for the expected texel-area as a function of plane parameters, camera parameters and texel location.

Drawing 1 illustrates the camera geometry and the symbols we use. We consider an image of a planar textured surface, using the pinhole camera model. Drawing 1 shows a slice that is perpendicular to the line of intersection of the image plane and the textured plane; the slice is distance y from the focal point. Both the image plane and the textured plane are perpendicular to the paper that the figure is drawn on.

Drawing 1 -- Scene layout and camera geometry



To simplify the derivation of the relationship between F_i and F_p we define two coordinate systems: a camera coordinate system and a plane coordinate system The orientation of both of these coordinate systems depends on the placement of the textured surface relative to the image plane (the x and u axes are chosen to align with the tilt direction). No loss of generality is involved: the coordinate systems are tools of the derivation and thus may be defined in any way we choose.

Here is a complete list of symbols used in this section. Many of these symbols are illustrated above in Drawing 1.

(x, y, z) A point in the camera coordinate system is denoted by (x, y, z). This is a left-handed coordinate system with the origin at the focal point. The view direction is along the positive z axis. The positive x axis points in the tilt direction, ie the direction of greatest depth increase.

(u, v, w) A point in the plane coordinate system is denoted by (u, v, w). This is a left-handed coordinate system with the origin at (0, 0, g) in camera coordinates. The u and v axes lie in the textured plane; thus the w component is zero for all points on the textured plane. The u axis is chosen so that after projection onto the image plane it becomes parallel with the x axis. Thus the v axis is parallel to the y axis.

From Drawing 1 we see how to convert the coordinates of a point on the textured plane to camera coordinates. A point on the textured plane, denoted by $(u, v, 0)$ in plane coordinates, is denoted by (x, y, z) in camera coordinates, where

$$x = u \cos S \qquad y = v \qquad z = u \sin S + g$$

S, T The slant and tilt of the textured plane are denoted by S and T respectively. (*Slant* and *tilt* are defined in Section 1.5.)

f, g The focal length is denoted by f and the distance along the optic axis from the focal point to the textured surface is denoted by g.

F_i, F_p, F_c The foreshortened dimension of a texel (texel extent measured in the direction of greatest depth increase; this is the minor axis length in Figure 1) is denoted by F_i for an image texel and by F_p for the physical texel. F_c denotes the value of F_i that would be measured for a texel located precisely at the image center.

U_i, U_p, U_c The unforeshortened dimension of a texel (texel extent measured in the direction of constant depth; this is the major axis length in Figure 1) is denoted by U_i for an image texel and by U_p for a physical texel. U_c denotes the value of U_i that would be measured for a texel located precisely at the image center.

A_i, A_c A_i denotes the area of a texel anywhere in the image. A_c denotes the value of A_i that would be measured for a texel located precisely at the image center.

θ The angle $\theta = \text{atan}(x_i / f)$ for an image point with coordinates (x_i, y, f). In order to compute θ for a given image location we need to know the tilt direction (since the orientation of the x axis depends on the tilt), as well as the the field-of-view of the camera lens (θ is larger in an image formed with a wide-angle lens than in an image formed with a telephoto lens).

6.2. The foreshortened texel dimension

We wish to find an expression for F_i, the observed length of the foreshortened texel dimension. Drawing 1 illustrates the derivation. The foreshortened texel dimension is parallel to the u axis. Thus the two endpoints of F_p are located at $(u_p, v, 0)$ and $(u_p', v, 0)$. In camera coordinates we denote these same two points by (x_p, y, z_p) and (x_p', y, z_p') respectively. F_i, the image extent corresponding to F_p has endpoints at (x_i, y, f) and (x_i', y, f). From the geometry in Drawing 1 we see that

$$\frac{x_i}{f} = \frac{x_p}{z_p} \qquad \frac{x_i'}{f} = \frac{x_p'}{z_p'}$$

Therefore

$$F_i = x_i' - x_i = f\left\{\frac{x_p'}{z_p'} - \frac{x_p}{z_p}\right\} = f\left\{\frac{u_p' \cos S}{z_p'} - \frac{u_p \cos S}{z_p}\right\}$$

$$= f \cos S \, \frac{u_p' z_p - u_p z_p'}{z_p \, z_p'} = f \cos S \, \frac{u_p'(g + u_p \sin S) - u_p(g + u_p' \sin S)}{z_p \, z_p'}$$

$$= f \cos S \, \frac{(u_p' - u_p)\, g}{z_p \, z_p'}$$

Since $F_p = u_p' - u_p$, we have

$$F_i = f \cos S \, \frac{F_p}{g} \, \frac{g^2}{z_p \, z_p'} \tag{6.1}$$

In order to simplify this expression, we derive an alternate expression for g / z_p:

$$g = z_p - u_p \sin S = z_p - x_p \tan S$$

$$\frac{g}{z_p} = 1 - \frac{x_p}{z_p}\tan S \quad = \quad 1 - \tan\theta \tan S \tag{6.2}$$

Substituting this expression for g/z_p (and a similar expression for g/z_p') into Equation (6.1), we obtain

$$F_i = F_p \frac{f}{g} \cos S \,(1 - \tan\theta \tan S)(1 - \tan\theta' \tan S)$$

If we make the approximation that the view angle does not change significantly across the texel, then θ is effectively constant across the texel, so $\theta = \theta'$. Then

$$F_i = F_p \frac{f}{g} \cos S \,(1 - \tan\theta \tan S)^2$$

We would like to convert this equation into a form that is independent of the focal length f and the surface depth g. Setting θ to zero we obtain an expression for F_c, the foreshortened dimension of a texel measured at the image center:

$$F_c = F_p \frac{f}{g} \cos S$$

Therefore the foreshortened dimension of a texel anywhere in the image is related to the foreshortened dimension of a texel at the image center by

$$F_i = F_c (1 - \tan\theta \tan S)^2 \tag{6.3}$$

As a reminder, the only approximation made in the above derivation is that the view angle does not change significantly across the texel, so that θ is effectively constant across the texel.

6.3. The unforeshortened texel dimension

The unforeshortened texel dimension lies along a line of equal distance from the image plane. Thus x and z are constant along the unforeshortened dimension of the physical texel; the endpoints of the physical texel are denoted by (x, y_p, z) and (x, y_p', z). We see that

$$U_p = y_p' - y_p \qquad \text{and} \qquad U_i = \frac{f}{z} y_p' - \frac{f}{z} y_p$$

so

$$U_i = U_p \frac{f}{z} \quad = \quad U_p \frac{f}{g} \frac{g}{z}$$

Substituting the expression for g/z from Equation (6.2), we obtain

$$U_i = U_p \frac{f}{g} (1 - \tan\theta \tan S)$$

Setting θ to zero we find the unforeshortened texel dimension at the image center:

$$U_c = U_p \frac{f}{g}$$

Therefore the unforeshortened dimension of a texel anywhere in the image is related to the unforeshortened dimension of a texel at the image center by

$$U_i = U_c (1 - \tan\theta \tan S) \tag{6.4}$$

6.4. The projected texel area

Assuming that the area A_i of an image texel is proportional to the product of F_i and U_i, we have $A_i = k\, F_i U_i$, where k is a constant of proportionality which depends upon the texel shape. Then, from Equations (6.3) and (6.4)

$$A_i = k\, F_c U_c (1 - \tan\theta\, \tan S)^3 \;=\; A_c (1 - \tan\theta\, \tan S)^3 \qquad (6.5)$$

The only approximation made in this derivation is that the view angle does not change significantly across the texel, so that θ is effectively constant across the texel.

Using Equation (6.5), we can predict the area of a texel located anywhere in an image of a textured planar surface. The following values are needed to make the prediction:

- A_c, the area that would be measured for a texel located at the center of the image.
- S and T, the slant and tilt of the textured plane.
- Field of view of the camera (the ratio of the film-width to the focal length). In order to calculate θ for a particular image location, we need the tilt of the textured plane as well as the field of view of the camera lens.

In our work we assume that the field of view of the camera lens is a known quantity. The other three quantities, (A_c, S, T), form the parameter space we search to find the best planar fit for a given texture image. This is discussed further in Section 7.

The Equations (6.3), (6.4) and (6.5) describe the appearance of texels in an image of a planar textured surface covered with identical texels. The texels are assumed to show no three-dimensional relief. Clearly, the textured surfaces typically occurring in natural scenes violate these assumptions. Sections 6. and 7. demonstrate that the equations are nevertheless useful for finding planar approximations to the textured surfaces visible in a variety of real images.

7. SURFACE ESTIMATION AND TEXEL IDENTIFICATION

Our goal in analyzing image texture is to find a spatial layout of homogeneously textured surfaces that could result in the given image texture. We do this by testing many spatial layouts and choosing the one that best matches the observed image texture.

A set of candidate texels is derived from the image using the methods described in Section 5. By finding a surface arrangement that is consistent with a maximal subset of the candidate texture elements, we calculate the surface parameters at the same time that we choose the true texels from among the candidates.

The current implementation is restricted to fitting a single planar surface to the image, based only on the observed areas of the candidate texture elements. The method could be extended to fit more complex surfaces, or to fit separate planar surfaces to different parts of the image. The method could also be extended to use additional properties of the candidate texture elements (aspect ratio, contrast, density) to obtain a more informed planar fit. Ideally, we would like to find a planar fit that is supported by parallel sources of information -- the observed changes (across the image) of texel area, aspect ratio, contrast and density can all give separate evidence to support the hypothesized surface arrangement. We have performed experiments with planar fits based on the aspect ratio gradient and the perspective gradient. (The perspective gradient is the gradient of unforeshortened texel extents; this is the gradient of major axis lengths in Figure 1). In the current implementation the extraction of elongated regions is not accurate; therefore these additional gradients do not provide much information beyond that obtained from the area gradients. An important step in extending this work is to develop a shape primitive that extracts and represents elongated regions more accurately than the disks do.

Having extracted candidate texels from an image of a textured surface, we find the orientation of the textured plane that best agrees with the observed areas of the candidate texels. A planar surface is characterized by the triple (A_c, S, T), where A_c is the texel area expected in the image center, S is the slant, and T is the tilt. In order to find the best planar fit for the image texture, we discretize the possible values of A_c, S and T, and evaluate the merits of each possible planar fit. For each choice of (A_c, S, T), the expected texel area is computed at each image location. These expected areas are compared to the region areas actually occurring in the image, and a fit-rating is computed for the plane. The plane that receives the highest fit-rating is selected as the estimate of the textured surface. The candidate texels that support the best planar fit are interpreted as true image texture elements (another planar fit may be performed for the left-over regions, to extract a second texture field).

For efficiency, the best planar fit is determined using a two-stage process. An initial coarse fit is performed using increments of 5° for slant, 10° for tilt, and 100% for A_c. The A_c values are chosen to increase exponentially because area-discrepancies are measured as a ratio of expected to actual areas. To refine the planar fit, a more detailed search of the (A_c, S, T) space is done in the neighborhood of the best plane from the coarse fit. Slant is stepped in increments of 2.5°, tilt is stepped in increments of 5°, and A_c is stepped in increments of less than 25%.

To evaluate a particular planar fit, the area of each potential texture element is compared with the texel area predicted by the parameters (A_c, S, T). The predicted texel area is computed using the following equation, which is derived in Section 6. (The angle θ depends on tilt and image location; the image

location used is the centroid of the potential texture element).

$$A_i = A_c (1 - \tan\theta \tan S)^3$$

If the expected and actual areas are similar, the candidate texel supports the planar fit well. The total support for each planar fit is:

$$\text{fit-rating} = \sum_{\textit{all regions}} (\text{region area}) \; |\text{region contrast}| \; e^{-(\text{region-fit})^2/4} \tag{7.1}$$

where

$$\text{region-fit} = \frac{\max(\text{expected area, actual area})}{\min(\text{expected area, actual area})}$$

The region-fit is 2.0 for a candidate texel that is either half as big or twice as big as the size predicted by the planar fit. As seen from Equation (7.1), the contribution made by a region falls off sharply as the region area deviates from the expected value. The regions make a contribution proportional to their area: this compensates for the fact that small regions outnumber large regions. Region contribution is proportional to the contrast of the region: higher contrast regions are perceptually more important and thus should have more influence on the planar fit. (We have tried performing planar fits where the "region contrast" term is left out of Equation (7.1). This works surprisingly well: the parameters of the best planar fit do not change much. However, the peak of the fit-rating values in (A_c, S, T) space is less pronounced.)

In Equation (7.1), the summation is written to be over all regions. This is not strictly true: in those image locations where multiple regions are possible (when a single disk participates in the formation of several candidate texture elements), the sum includes only the candidate texel whose area best agrees with the hypothesized planar fit.

Results obtained for a variety of images are illustrated in Figures 5 to 38. Parts (c) and (d) of each figure show the texels that are extracted on the basis of the best planar fit. Part (e) of each figure is a synthetic image illustrating the (A_c, S, T) parameters of the best planar fit. The height fields in part (f) of each figure show fit-rating as a function of slant and tilt, with A_c fixed at the value that produces the best planar fit for the texture in question. The height fields flatten out near the back because tilt becomes less important as slant decreases; the planar fit is independent of tilt when the slant is zero. The graphs in part (g) of each figure show fit-rating as a function of A_c, with slant and tilt fixed at the values that produce the best planar fit for the texture in question. The fit-rating values change smoothly as a function of A_c, slant and tilt. The absence of secondary peaks and ridges makes it easy to identify the best planar fit. These results are discussed further in Section 8.

8. IMPLEMENTATION-SUMMARY AND RESULTS

Some implementation details for region detection and surface fitting were given in Sections 5 and 7 respectively. In this section we present a summary of the implementation and the results obtained on natural images. Results are shown for a number of textures, so that the strengths, weaknesses and generality of the implementation may be judged. All of the images are processed the same way; the method has no parameters that need to be tuned to particular images.

8.1. Summary of the implementation

Here we list the processing steps used on all of the images used in our experiments. The processing of an image I is divided into three main phases: fit disks to the uniform image regions (Section 5.8.1.), construct potential texture elements from the disks (Section 5.8.2.), and fit a planar surface to the candidate texels (Section 7.2.).

Fit disks to the uniform image regions

(1) Compute the convolutions $\nabla^2 G*I$ and $\frac{\partial}{\partial\sigma}\nabla^2 G*I$ for the following six σ values: $\sqrt{2}$, $2\sqrt{2}$, $3\sqrt{2}$, $4\sqrt{2}$, $5\sqrt{2}$ and $6\sqrt{2}$. (The center lobes of the six $\nabla^2 G$ filters have diameters of 4, 8, 12, 16, 20 and 24 pixels respectively.)

(2) Mark the locations where disks will be fit. To analyze the positive-contrast regions of the original image, mark all local maxima in the $\nabla^2 G*I$ images. To analyze the negative-contrast regions of the original image, mark all local minima in the $\nabla^2 G*I$ images.

(3) At each marked location, use the measured $\nabla^2 G*I$ and $\frac{\partial}{\partial\sigma}\nabla^2 G*I$ values to compute a disk diameter and disk contrast:

$$D = 2\sigma\sqrt{\sigma(\tfrac{\partial}{\partial\sigma}\nabla^2 G*I)/(\nabla^2 G*I) + 2} \qquad C = \frac{2\sigma^2}{\pi D^2}\, e^{D^2/8\sigma^2}\,(\nabla^2 G*I)$$

Retain only the disks where $w-2 \le D \le w+2$ (w is the width in pixels of the center lobe of the $\nabla^2 G$ filter).

Construct potential texture elements from the disks

To form the list of potential texture elements, extract all subsets of disks that are spatially connected and contain no concavities greater than 90°. If a concavity is in the range 50° to 90°, use the disks to form three potential texture elements: one large region consisting of all the disks, and two smaller regions resulting from splitting the large region at the concavity[1]. Mark mutual exclusion between potential texture elements that share a disk: at most one of them can contribute support to a planar fit and be chosen as a true texture element.

Fit a planar surface to the candidate texels

A_c is the texel area expected in the image center, S is the slant, and T is the tilt of a hypothesized

[1]Region splitting is implemented as follows. We begin with a set P of overlapping disks, which together cover an image region R. The largest concavity in R is found by computing the angles formed by every pair of neighboring disks on the border of R. Suppose that X and Y are two neighboring disks on the border of R, and that they form a concavity that should cause a split into smaller, more convex regions. The concavity is split by (1) removing X from P and repeating the above process, and then (2) removing Y from P and repeating the above process.

planar fit. For a coarse fit, choose A_c from the set {10, 20, 40, 80, 160, 320, 640}, choose S from {0°, 5°, 10°, ..., 70°, 75°, 80°}, and choose T from {0°, 20°, 40°, ..., 300°, 320°, 340°}. To perform a fine fit in the neighborhood of the best plane from the coarse fit, change S in increments of 2.5°, T in increments of 5°, and A_c in increments of less than 25%.

The expected texel area for a particular choice of (A_c, S, T) is computed as $A_i = A_c(1 - \tan\theta \tan S)^3$. (See Section 6 for a definition of the angle θ.) Evaluate a planar fit by adding contributions from each potential texel:

$$\text{fit-rating} = \sum_{\textit{all regions}} (\text{region area}) \; |\text{region contrast}| \; e^{-(\text{region-fit})^2/4}$$

where $\text{region-fit} = \dfrac{\max(\text{expected area, actual area})}{\min(\text{expected area, actual area})}$. Select the plane that receives the highest fit-rating as the best estimate of the textured surface. Identify texture elements as those regions that have an area close to the area expected by the best planar fit.

8.2. The images

Parts (a) of Figures 5 to 38 show images of seventeen natural textures. A few of the images are photographs of outdoor scenes in Urbana, Illinois. The rest are illustrations in books which have been rephotographed. All of these images are digitized off of the photographic negatives using a drum scanner. The images are 512 by 512 pixels; the image sizes in the figures vary because image borders have been trimmed. Table 2 indicates the source of each image.

TABLE 2

Description		Source of image
A rock pile	Figures 5 and 6	Outdoor scene in Urbana, Illinois
An aerial view of houses	Figures 7 and 8	Silverman [1983], page 221
Snow Geese flying over water	Figures 9 and 10	Bourke-White [1972], page 201
Muslims at a mosque	Figures 11 and 12	Bourke-White [1972], page 168
Fleecy clouds	Figures 13 and 14	Strache [1956], plate 5
Audience at a 3D movie	Figures 15 and 16	Life [1984], plate 1
Sunflowers	Figures 17 and 18	Landscape [1984], page 75
A tree trunk	Figures 19 and 20	Outdoor scene in Urbana, Illinois
Bathers on the Ganges	Figures 21 and 22	Adams and Newhall [1960], page 42
A plowed field	Figures 23 and 24	Bourke-White [1972], page 185
A field of flowers	Figures 25 and 26	Gullers and Strandell [1977], page 5
Water lillies	Figures 27 and 28	Thomas [1976], page 97
Ripple marks in a shallow sea	Figures 29 and 30	Strache [1956], plate 14
Water Hyacinths	Figures 31 and 32	Thomas [1976], page 14
The Toulumne River	Figures 33 and 34	Adams and Newhall [1960], page 64
Sand by the Adriatic Sea	Figures 35 and 36	Landscape [1984], page 95
Fallen leaves	Figures 37 and 38	Outdoor scene in Urbana, Illinois

8.3. Discussion of the results

Figures 5 to 38 illustrate the results we obtain on the seventeen images of natural textures. The results obtained for each image are illustrated in two successive figures. The first figure shows the results obtained on the positive-contrast image regions, whereas the second figure shows the results obtained on the negative-contrast image regions.

The original image is shown in part (a) of Figures 5 to 38. Part (b) of each figure shows the disks that model the regions of uniform gray level in the original image. It is impossible to display all the disks in a single image, since many disks are spatially contained in larger disks. This spatial containment typically means that either (1) the large disk is part of a texture element and the small disks are subtexture, or (2) the small disks are texture elements and the large disk is supertexture. In case (1) the large disk usually has higher contrast than the smaller disks, whereas in case (2) the smaller disks usually have higher contrast than the large disk. Wherever disks overlap, our figures shows the disk of higher contrast. Therefore most subtexture disks in part (b) of Figures 5 to 38 are not visible: they are covered by a larger, higher-contrast disk corresponding to part of a texture element. Refer to Figure 4 for an illustration of the complete set of disks found for a particular image (the rock pile).

The detected texels are shown in parts (c) and (d) of Figures 5 to 38: these are all image regions having area within a factor of two of the area expected by the best planar fit. The parameters of the best planar fit are illustrated by the synthetic texture images in part (e) of each figure.

Parts (f) and (g) of Figures 5 to 38 illustrate the change of fit-rating as a function of A_c, slant and tilt. The height fields in part (f) of each figure show fit-rating as a function of slant and tilt, with A_c fixed at the value that produces the best planar fit for the texture in question. The height fields flatten out near the back because tilt becomes less important as slant decreases; the planar fit is independent of tilt when the slant is zero. The graphs in part (g) of each figure show fit-rating as a function of A_c, with slant and tilt fixed at the values that produce the best planar fit for the texture in question.

The shape of the fit-rating peak is related to the properties of the image texture. A sharp fit-rating peak indicates that the texels have small size variance. This is illustrated by the aerial view of houses (Figures 7 and 8) and by the field of sunflowers (Figures 17 and 18). If the texel sizes have larger variance, as for the clouds (Figures 13 and 14) and the rock pile (Figures 5 and 6), then the peak is much broader. (In the rock-pile image, the non-planarity of the original textured surface also contributes to the broadness of the fit-rating peak.) The texels shown in parts (c) and (d) of the figures are those candidate texels having area within a factor of two of the area expected by the planar fit. Using this same factor of two for all images causes incomplete extraction of texels in images where texel size is highly variable. More complete texel extraction can be achieved by adjusting the criteria for choosing texels from the set of candidate texels: the criteria should vary as a function of the broadness of the fit-rating peak in (A_c, S, T) space.

The accuracy of the results may be illustrated in two ways. Firstly, the reader can compare his perception of the textured surfaces (part (a) of Figures 5 to 38) with the planar surface fitted by the program (part (e) of Figures 5 to 38). Agreement with human perception is quite good for many of the images. Secondly, since the processing of the positive-contrast and negative-contrast regions is performed totally independently, the agreement between the slants and tilts obtained by the two analyses strengthens the confidence in the results. (The A_c parameters are not expected to be similar for the positive-contrast and negative-contrast regions -- the positive-contrast and negative-contrast regions may be of very different sizes.) However, the two analyses may not always lead to the same estimates of slant and tilt, because a texture may not be homogeneous in both texel size and texel separation. Thus, an agreement among multiple analyses (such as the two discussed here) must not be required. A method of selecting and integrating

the pertinent analyses in a given case must be devised. Such inferencing from gradients of multiple texture properties has not been addressed in the work reported in this paper.

Table 3 summarizes the planar fits obtained for all images. These fits use slants that are multiples of 2.5° and tilts that are multiples of 5°. The slant and tilt values computed from the positive-contrast and negative-contrast regions are often within 10° of each other. Seven of the 17 images have differences less than 10°; nine of the images have differences less than 15°. For reference, a 30° difference in tilt is equal to the angular distance between adjacent numbers on a clock face. A 30° difference in slant, on the other hand, is a more serious error. In many of those images that have a large discrepancy between the two planar fits, attributes of the original texture lead us to expect the fits to differ in accuracy. We have identified four reasons for the observed discrepancies. In the field of flowers (Figure 25) and the water lillies (Figure 27), the spaces between the texels are less regular than are the areas of the texels; therefore the fit to the negative-contrast regions is not as accurate as the fit to the positive-contrast regions. A second reason the background regions produce inaccurate results is because the properties of the physical texels are more important than the properties of background regions. In images where the physical texels are separated by gaps, the linear distance between image texels carries more information than does the shape or area of the background regions. Thus, the results for the negative-contrast regions of the movie image (Figure 16) and the lilly pad image (Figure 28) are inaccurate because the area of the background regions poorly reflects the inter-texel spacing. A third reason for discrepancies between the two slant and tilt estimates is a large variability in texel area (as occurs in Figure 11, the image of Muslims at a mosque). This causes a broad peak in the planar fit space (part (f) of Figure 11); hence the exact peak location is not as accurate for these images as for others. A fourth reason for inaccurate results is that the current extraction

TABLE 3

Description	Figures	Fit to positive-contrast regions			Fit to negative-contrast regions			Difference	
		A_c	slant	tilt	A_c	slant	tilt	slant	tilt
A rock pile	5, 6	40	62.5°	65°	40	60°	75°	2.5°	10°
Aerial view of houses	7, 8	35	62.5°	95°	60	67.5°	110°	5°	15°
Birds flying over water	9, 10	35	45°	80°	40	57.5°	100°	12.5°	20°
Muslims at a mosque	11, 12	160	27.5°	50°	120	42.5°	100°	15°	50°
Fleecy clouds	13, 14	100	55°	275°	160	55°	280°	0°	5°
3D movie audience	15, 16	280	45°	105°	320	7.5°	330°	large	
Sunflowers	17, 18	160	70°	95°	200	70°	90°	0°	5°
A tree trunk	19, 20	70	65°	345°	80	42.5°	0°	25.5°	15°
Bathers on the Ganges	21, 22	100	45°	80°	80	65°	85°	20°	5°
A plowed field	23, 24	80	42.5°	40°	100	65°	80°	22.5°	40°
A field of flowers	25, 26	50	70°	90°	140	52.5°	20°	large	
Water lillies	27, 28	120	75°	90°	160	52.5°	70°	22.5°	20°
Ripples	29, 30	50	52.5°	105°	120	62.5°	105°	10°	0°
Water Hyacinths	31, 32	100	37.5°	80°	100	40°	80°	2.5°	0°
The Toulumne River	33, 34	25	57.5°	85°	40	65°	95°	7.5°	10°
Sand	35, 36	240	40°	80°	200	55°	80°	15°	0°
Fallen leaves	37, 38	40	60°	90°	50	62.5°	95°	2.5°	5°

of uniform regions fragments non-compact regions in an arbitrary way, increasing the variabilities of the measured areas. This effect can be seen in the background of the movie image (Figure 16).

For nearly all of the images, at least one of the two analyses produces results that are in good agreement with human perception. Future work may produce a method for automatically determining which analysis -- the analysis of positive-contrast regions, or the analysis of negative-contrast regions -- has produced the most accurate results.

9. SUMMARY AND CONCLUSIONS

We have presented a general discussion of the problem of recovering scene-layout information from the texture cues present in an image. We argue that extraction of texels is useful and perhaps even necessary for correct interpretation of texture gradients in the face of subtexture, multiple texture fields, and occlusions. In order to separate texture elements from other regions (such as subtexture regions or texels from a second texture field) it is necessary to integrate the processes of texel identification and surface estimation. The processing of a texture image should ideally be an integrated analysis of all relevant texture gradients, including area gradients, aspect-ratio gradients and density gradients.

We have presented an implementation that is based on these ideas; the implementation is restricted to the detection of gradients of texel area. We derived a region detector based on the response of an ideal disk to convolution with a Laplacian-of-Gaussian ($\nabla^2 G$) over a range of scales. The output of the region detector is used to form a list of candidate texels. These candidate texels then provide the evidence needed to choose a good planar fit to the image texture; at the same time, the best planar fit is used to choose the true texels from among the candidates. Both positive-contrast and negative-contrast image regions are analyzed for texture information. Results are shown for a wide variety of natural textures.

The region detector and the techniques used to derive it may prove useful in computer vision applications other than texture analysis. The extraction of texture elements, especially of elongated texture elements, needs to be improved. We are interested in the development of an elongated shape primitive to complement or replace the circular disk primitive obtained from the $\nabla^2 G$ scale-space. We do not have to restrict our attention to the $\nabla^2 G$ filter; other filters may be more amenable to analysis. A better treatment of elongated texels will allow additional texture gradients, such as gradients of aspect ratio, to be measured. If several texture gradients are analyzed, methods must be developed to combine the information obtained from each gradient. As we have discussed in Section 2.3., the relative accuracy of the various texture gradients varies from image to image. In combining the results from separate analyses of several texture gradients, it is important to determine which of the texture gradients have given the most accurate results.

Our current implementation produces a planar approximation to the textured surface seen in an image. Better shape approximations for the textured surface could be obtained in various ways. Planar surface patches could be fit to subwindows of the image. However, the choice of window-sizes is a difficult problem. The texture data may be too variable to permit accurate fitting of small planar patches; a method is needed to judge when a planar patch is large enough to allow an accurate estimate of slant and tilt. It may be possible to recognize locations of texture curvature directly, by looking for changes in the compression gradient. As discussed in Section 2.4.2., size and density gradients are important in judging the slant of flat surfaces, whereas the compression gradient is the most important gradient for perception of curved surfaces. Distance and foreshortening effects cause texture features to vary gradually across the image, except at discontinuities of depth or surface orientation, and at boundaries between different surface textures. Methods for recognizing and locating these discontinuities are needed.

Analysis of the relationships between various texture fields could lead to a better understanding of the physical structure of the texture. For example, one could note the relationship between the houses and their shadows in Figure 7, between the heads and the facial features in Figure 15, and between the centers of the flowers and the petals in Figure 17. Currently, we treat these various components of the physical texture elements as separate texture fields, without noting the systematic relationships among them.

A significant aspect of this work is that it has been tested on enough real images to demonstrate its strong and weak points. Unfortunately, texture is a fairly ill-defined concept. It is difficult to be rigorous with this subject, to give a precise definition of the problem and to list criteria for judging when the problem is solved. This paper has developed a method of texture analysis that passes the only test we have: it works fairly successfully on a wide range of images.

APPENDIX

THE $\nabla^2 G$ RESPONSES OF DISK AND BAR IMAGES

In this appendix we derive closed form expressions for the response of bar and disk images to a $\nabla^2 G$ filter. The symbol definitions of Section 5.2. are used.

Given a function $I(x,y)$ which describes the intensity of an image at (x,y), the $\nabla^2 G$ response of this image at (x,y) is given by the following convolution:

$$\nabla^2 G(x,y) \;*\; I(x,y) = \int_{-\infty}^{+\infty}\!\!\int \nabla^2 G(u,v)\, I(x-u, y-v)\, du\, dv$$

$$= \int_{-\infty}^{+\infty}\!\!\int \frac{2\sigma^2-(u^2+v^2)}{\sigma^4}\, e^{-(u^2+v^2)/2\sigma^2} I(x-u, y-v)\, du\, dv \qquad \text{(A.1)}$$

The class of functions $I(x,y)$ which have a closed-form solution for the integral of Equation (A.1) is quite limited. A closed form for $\int e^{u^2} du$ exists only when the bounds of integration are zero or infinite. Therefore we have a closed-form solution for the $\nabla^2 G$ response of an infinitely long bar, but cannot find a solution for a rectangle or one-way infinite bar.

A.1. Some useful integerals

We begin with a list of integrals which will be used in later derivations. The well-known identity

$$\int_{-\infty}^{+\infty} e^{-z^2} dz = \sqrt{\pi}$$

may be put into a slightly different form using the substitution $z = t/\sqrt{2}\sigma$:

$$\int_{-\infty}^{+\infty} e^{-t^2/2\sigma^2} dt = \sqrt{2\pi}\,\sigma \qquad \text{(A.2)}$$

Integrating by parts and using Equation (A.2) we have

$$\int_{-\infty}^{+\infty} t^2 e^{-t^2/2\sigma^2} dt = \sqrt{2\pi}\,\sigma^3 \qquad \text{(A.3)}$$

The *error function* is defined here as

$$erf\,(k) = \int_{-\infty}^{k} \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2} dx$$

Thus

$$\int_{-\infty}^{M} t^2 e^{-t^2/2} dt \;=\; -t\, e^{-t^2/2}\Big]_{-\infty}^{M} + \int_{-\infty}^{M} e^{-t^2/2} dt \;=\; -M\, e^{-M^2/2} + \sqrt{2\pi}\, erf\,(M) \qquad \text{(A.4)}$$

We use the following equation to solve for the $\nabla^2 G$ response of a disk image.

$$\int_0^{D/2} (2\sigma^2-\rho^2)\, e^{-\rho^2/2\sigma^2}\, \rho\, d\rho \;=\; \frac{D^2\sigma^2}{4}\, e^{-D^2/8\sigma^2} \tag{A.5}$$

In deriving this equation, we use integration by parts on the second integral ($\int u\, dv = uv - \int v\, du$, with $u=\rho^2$, $du = 2\rho d\rho$, $v = \sigma^2 e^{-\rho^2/2\sigma^2}$, and $dv = -\rho\, e^{-\rho^2/2\sigma^2}$):

$$\int_0^{D/2} 2\sigma^2\rho\, e^{-\rho^2/2\sigma^2} d\rho \quad + \quad \int_0^{D/2} -\rho^3\, e^{-\rho^2/2\sigma^2} d\rho$$

$$= \int_0^{D/2} 2\sigma^2\rho\, e^{-\rho^2/2\sigma^2} d\rho \quad + \quad \rho^2\sigma^2 e^{-\rho^2/2\sigma^2}\Big]_0^{D/2} \quad - \quad \int_0^{D/2} 2\rho\sigma^2 e^{-\rho^2/2\sigma^2} d\rho$$

$$= \rho^2\sigma^2 e^{-\rho^2/2\sigma^2}\Big]_0^{D/2} \quad = \quad \frac{D^2\sigma^2}{4}\, e^{-D^2/8\sigma^2}$$

A.2. $\nabla^2 G$ response of a step-edge image

The $\nabla^2 G$ response of a step-edge image will later be used to compute the $\nabla^2 G$ response of a bar image. Consider an image of a vertical step edge at $x=B$ defined by

$$\text{step-edge image:} \quad I(x,y) = \begin{cases} C & \text{if } x \geq B \\ 0 & \textit{elsewhere} \end{cases} \tag{A.6}$$

Using this definition of $I(x,y)$ in Equation (A.1) gives

$$\int_{-\infty}^{x-B} \int_{-\infty}^{+\infty} \frac{2\sigma^2-(u^2+v^2)}{\sigma^4}\, e^{-(u^2+v^2)/2\sigma^2} C\, dv\, du$$

$$= \frac{2C}{\sigma^2} \int_{-\infty}^{x-B} e^{-u^2/2\sigma^2} \int_{-\infty}^{+\infty} \left\{ (1-\frac{u^2}{2\sigma^2})\, e^{-v^2/2\sigma^2} - \frac{v^2}{2\sigma^2}\, e^{-v^2/2\sigma^2} \right\} dv\, du$$

Using the integrals given in Equations (A.2) and (A.3), this simplifies to

$$= \frac{2C}{\sigma^2} \int_{-\infty}^{x-B} e^{-u^2/2\sigma^2} \left\{ (1-\frac{u^2}{2\sigma^2})\sqrt{2\pi}\,\sigma - \frac{1}{2\sigma^2}\sqrt{2\pi}\,\sigma^3 \right\} du$$

$$= \sqrt{2\pi}\,\sigma \frac{2C}{\sigma^2} \int_{-\infty}^{x-B} \left\{ 1 - \frac{u^2}{2\sigma^2} - \frac{1}{2} \right\} e^{-u^2/2\sigma^2} du$$

$$= \frac{\sqrt{2\pi}\,C}{\sigma} \int_{-\infty}^{x-B} \left\{ 1 - \frac{u^2}{\sigma^2} \right\} e^{-u^2/2\sigma^2} du$$

Make the substitution $t = u/\sigma$ $(du = \sigma\, dt)$:

$$= \frac{\sqrt{2\pi}\,C}{\sigma} \int_{-\infty}^{\frac{x-B}{\sigma}} (1-t^2)\, e^{-t^2/2}\, \sigma\, dt$$

$$= 2\pi C \; erf\left(\frac{x-B}{\sigma}\right) - \sqrt{2\pi} C \int_{-\infty}^{\frac{x-B}{\sigma}} t^2 e^{-t^2/2} dt$$

Substituting Equation (A.4),

$$= 2\pi C \; erf\left(\frac{x-B}{\sigma}\right) - \sqrt{2\pi} C \left\{ -\frac{x-B}{\sigma} \; e^{-(x-B)^2/2\sigma^2} + \sqrt{2\pi}\, erf\left(\frac{x-B}{\sigma}\right) \right\}$$

Since the two *erf* terms cancel, we have

$$\nabla^2 G \text{ response of step edge:} \quad \frac{\sqrt{2\pi} C \, (x-B)}{\sigma} \, e^{-(x-B)^2/2\sigma^2} \tag{A.7}$$

This expression is the closed-form solution for the $\nabla^2 G$ response of a vertical step edge of intensity C located at $x=B$. This equation illustrates the zero-crossing of the Marr and Hildreth edge operator (Marr [1980]): for $x<B$ the response is negative, at $x=B$ the response is zero, and for $x>B$ the response is positive.

A.3. ∇^2G response of a bar image

Consider an image consisting of a bar of width B and intensity C on a zero background:

$$\text{bar image:} \quad I(x,y) = \begin{cases} C & \text{if } 0 \le x \le B \\ 0 & elsewhere \end{cases} \tag{A.8}$$

In order to compute the $\nabla^2 G$ response of such a bar, we take the sum of two step-edge responses. Using Equation (A.7), we have

$\nabla^2 G$ response of a bar = (response of a step up at $x = 0$) + (response of a step down at $x = B$)

$$= \sqrt{2\pi} \frac{C}{\sigma} \left\{ x \, e^{-x^2/2\sigma^2} - (x-B) \, e^{-(x-B)^2/2\sigma^2} \right\} \tag{A.9}$$

Substituting $x = B/2$ into Equation (A.9),

$$\nabla^2 G \text{ response at the center of a bar:} \quad \frac{\sqrt{2\pi} CB}{\sigma} \, e^{-B^2/8\sigma^2} \tag{A.10}$$

Taking the derivative with respect to sigma,

$$\frac{\partial}{\partial \sigma} \nabla^2 G \text{ response at the center of a bar:} \quad \sqrt{2\pi} CB \left\{ \frac{B^2}{4\sigma^4} - \frac{1}{\sigma^2} \right\} e^{-B^2/8\sigma^2} \tag{A.11}$$

The $\nabla^2 G_n$ response of a bar differs from the $\nabla^2 G$ response by a factor of $2\pi\sigma^2$. Thus we calculate

$$\nabla^2 G_n \text{ response at the center of a bar:} \quad \frac{CB}{\sqrt{2\pi}\,\sigma^3} \, e^{-B^2/8\sigma^2} \tag{A.12}$$

and

$$\frac{\partial}{\partial \sigma} \nabla^2 G_n \text{ response at the center of a bar:} \quad \frac{CB}{\sqrt{2\pi}} \left\{ \frac{B^2}{4\sigma^6} - \frac{3}{\sigma^4} \right\} e^{-B^2/8\sigma^2} \tag{A.13}$$

A.4. $\nabla^2 G$ response of a disk image

Consider an image consisting of a circular disk of diameter D and intensity C on a zero background:

$$\text{disk image:} \quad I(x,y) = \begin{cases} C & \text{if } x^2+y^2 \le D^2/4 \\ 0 & \textit{elsewhere} \end{cases} \tag{A.14}$$

We have not succeeded in finding a general closed-form solution for Equation (A.1) using this definition of $I(x,y)$. However, we can solve Equation (A.1) when $x=y=0$, giving the $\nabla^2 G$ response at the center of the disk. With x and y zero, Equation (A.1) becomes

$$\int_{-\infty}^{+\infty}\int \frac{2\sigma^2-(u^2+v^2)}{\sigma^4}\, e^{-(u^2+v^2)/2\sigma^2} I(-u,-v)\, du\, dv$$

Change to polar coordinates ($\rho^2 = u^2+v^2$; $du\, dv = \rho\, d\rho\, d\theta$) and using $I(x,y)$ from Equation (A.14),

$$= \frac{C}{\sigma^4} \int_{-\pi}^{\pi} \int_{0}^{D/2} (2\sigma^2-\rho^2)\, e^{-\rho^2/2\sigma^2}\, \rho\, d\rho\, d\theta$$

Substituting the solution to the inner integral from Equation (A.5),

$$= \frac{CD^2}{4\sigma^2}\, e^{-D^2/8\sigma^2} \int_{-\pi}^{\pi} d\theta$$

Thus

$$\nabla^2 G \text{ response at the center of a disk:} \quad \frac{\pi CD^2}{2\sigma^2}\, e^{-D^2/8\sigma^2} \tag{A.15}$$

Taking the derivative with respect to sigma,

$$\frac{\partial}{\partial\sigma}\nabla^2 G \text{ response at the center of a disk:} \quad \frac{\pi CD^2}{2}\left\{\frac{D^2}{4\sigma^5} - \frac{2}{\sigma^3}\right\} e^{-D^2/8\sigma^2} \tag{A.16}$$

The $\nabla^2 G_n$ response at the disk center differs from the $\nabla^2 G$ response by a factor of $2\pi\sigma^2$. Thus

$$\nabla^2 G_n \text{ response at the center of a disk:} \quad \frac{CD^2}{4\sigma^4}\, e^{-D^2/8\sigma^2} \tag{A.17}$$

and

$$\frac{\partial}{\partial\sigma}\nabla^2 G_n \text{ response at the center of a disk:} \quad \frac{CD^2}{4}\left\{\frac{D^2}{4\sigma^7} - \frac{4}{\sigma^5}\right\} e^{-D^2/8\sigma^2} \tag{A.18}$$

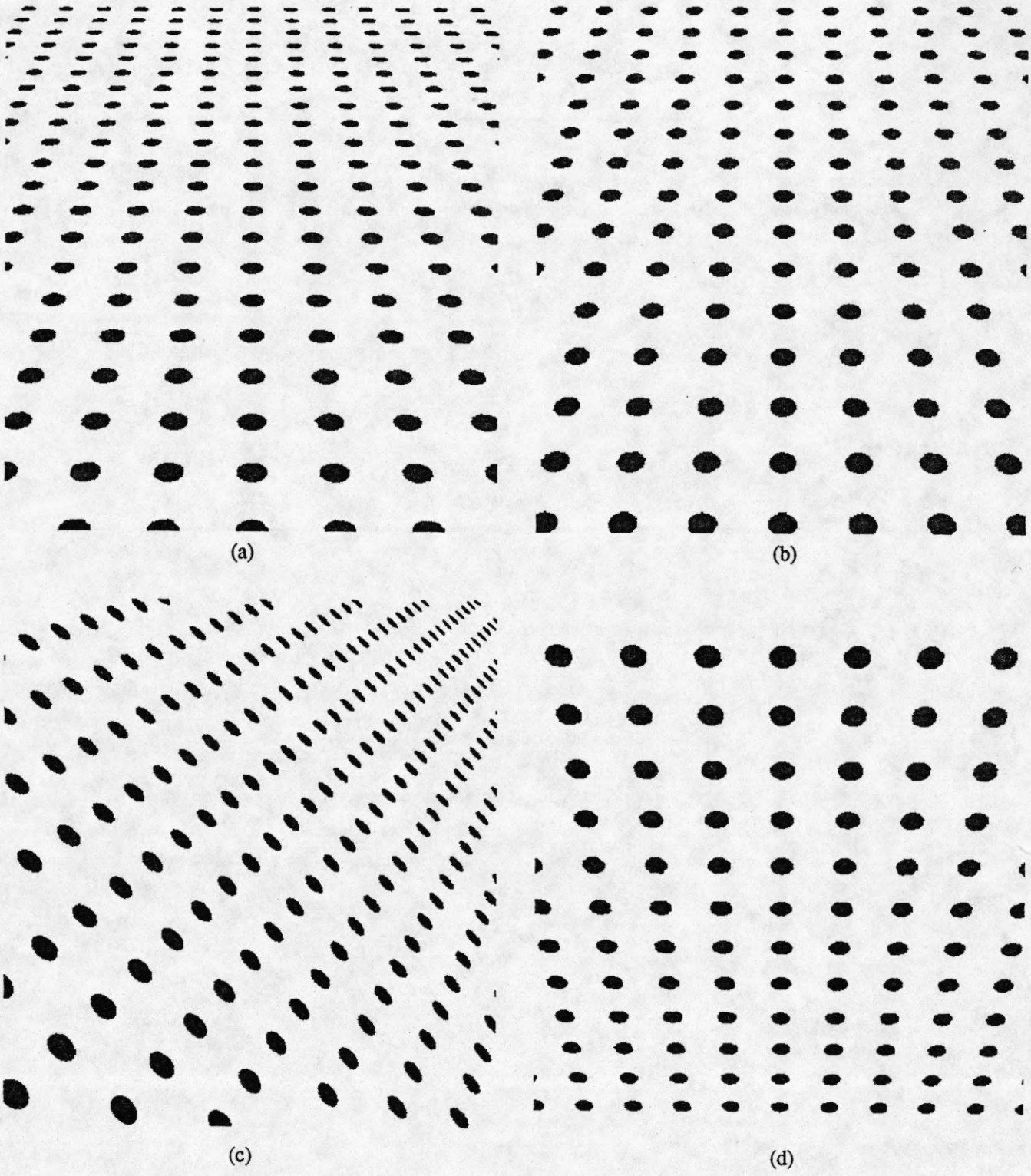


Figure 1

Synthetic textures illustrating various slants and tilts. Slant is the angle between the textured surface and the image plane. Tilt is the direction in which the surface normal projects in the image. (a) Slant 60°, tilt 90°. (b) Slant 50°, tilt 90°. (c) Slant 60°, tilt 45°. (d) Slant 45°, tilt 270°.

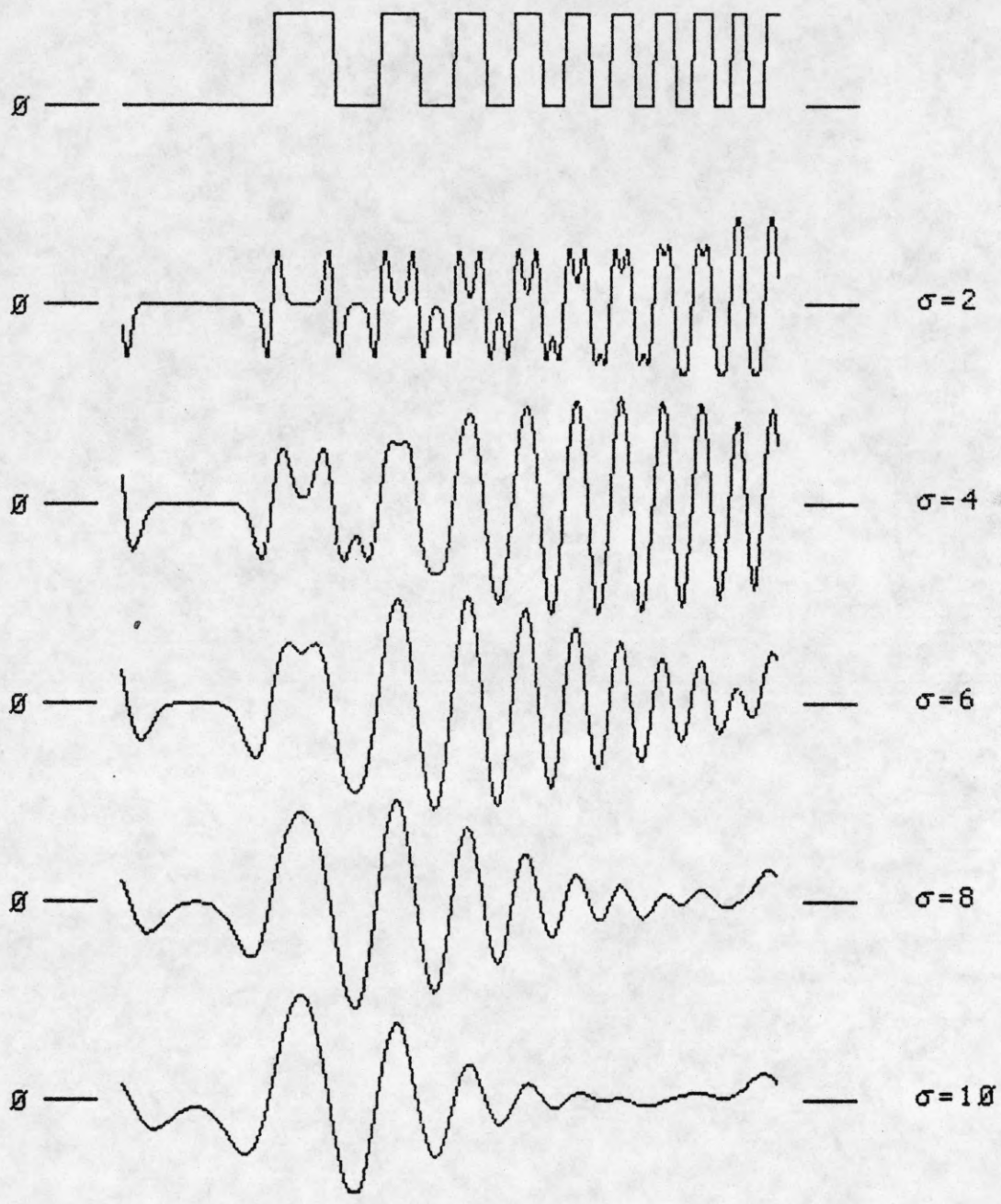
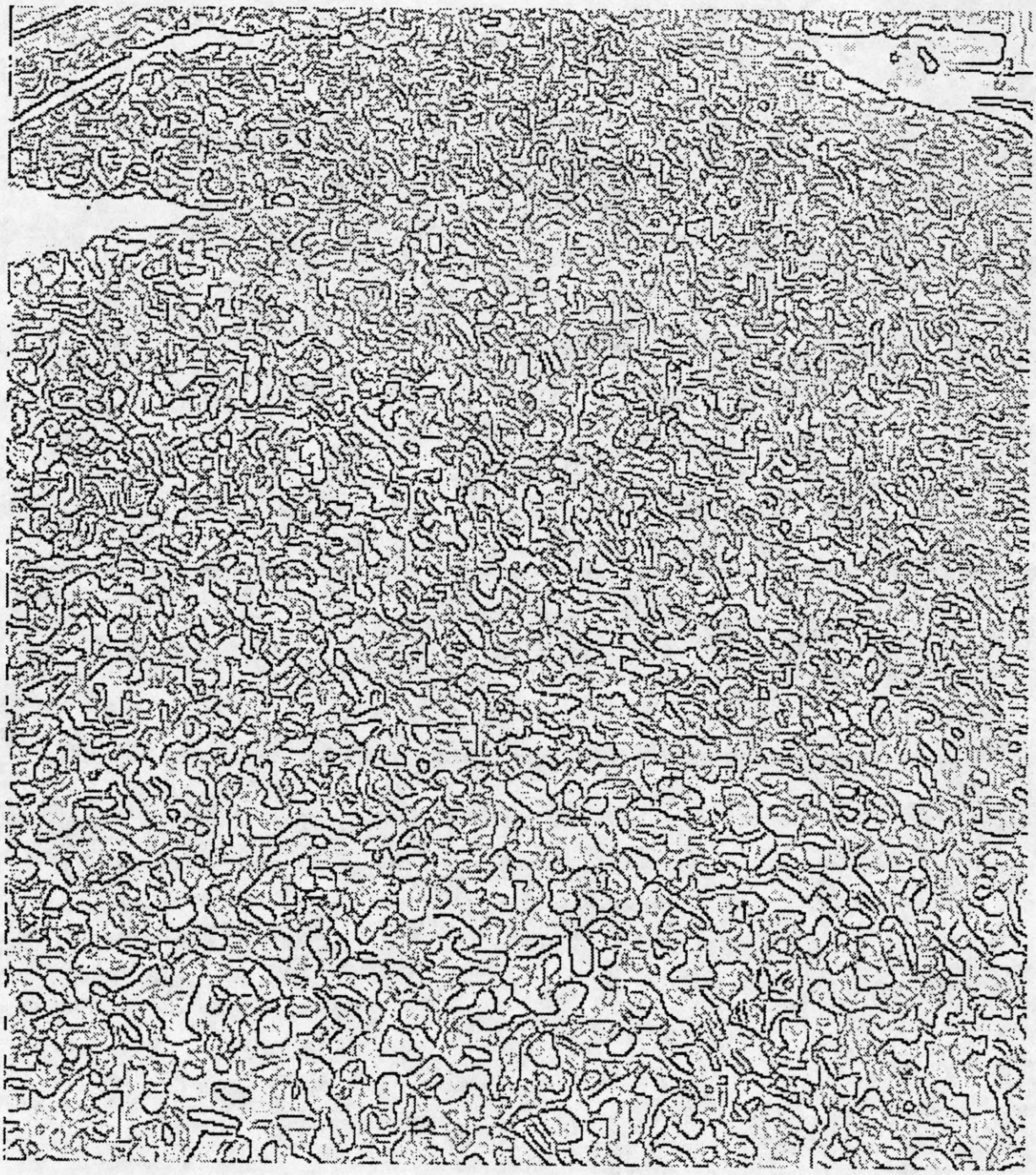


Figure 2

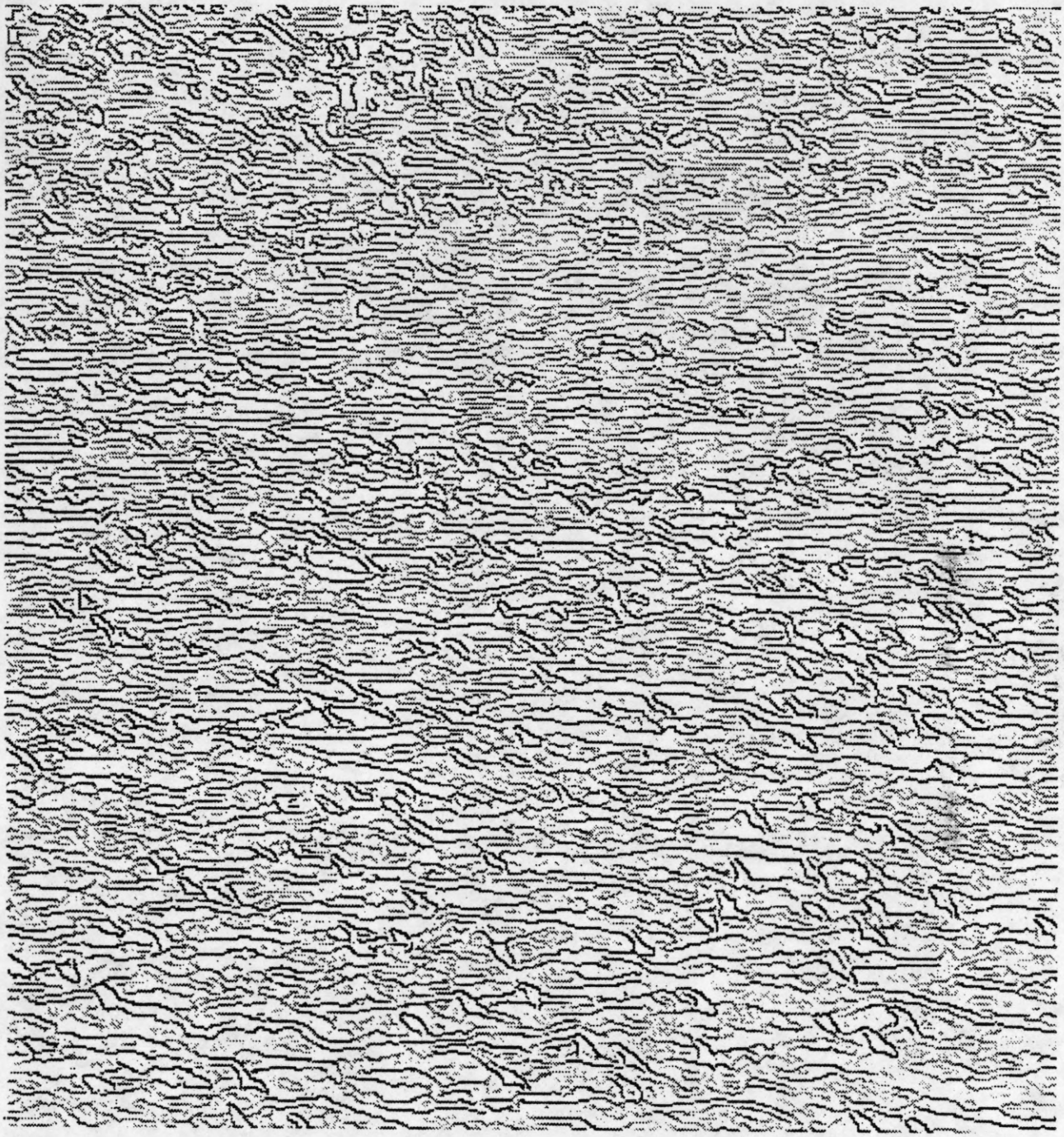
The top plot is a cross section of an image of varying-width bars. Subsequent plots, all on the same vertical scale, show the result of convolving the image with $\nabla^2 G$ filters of various sizes. The impulse response of the one-dimensional $\nabla^2 G$ filter is $(\sqrt{2\pi}/\sigma^3)(\sigma^2 - x^2)\, e^{-x^2/2\sigma^2}$. A circular convolution is used.



(a)

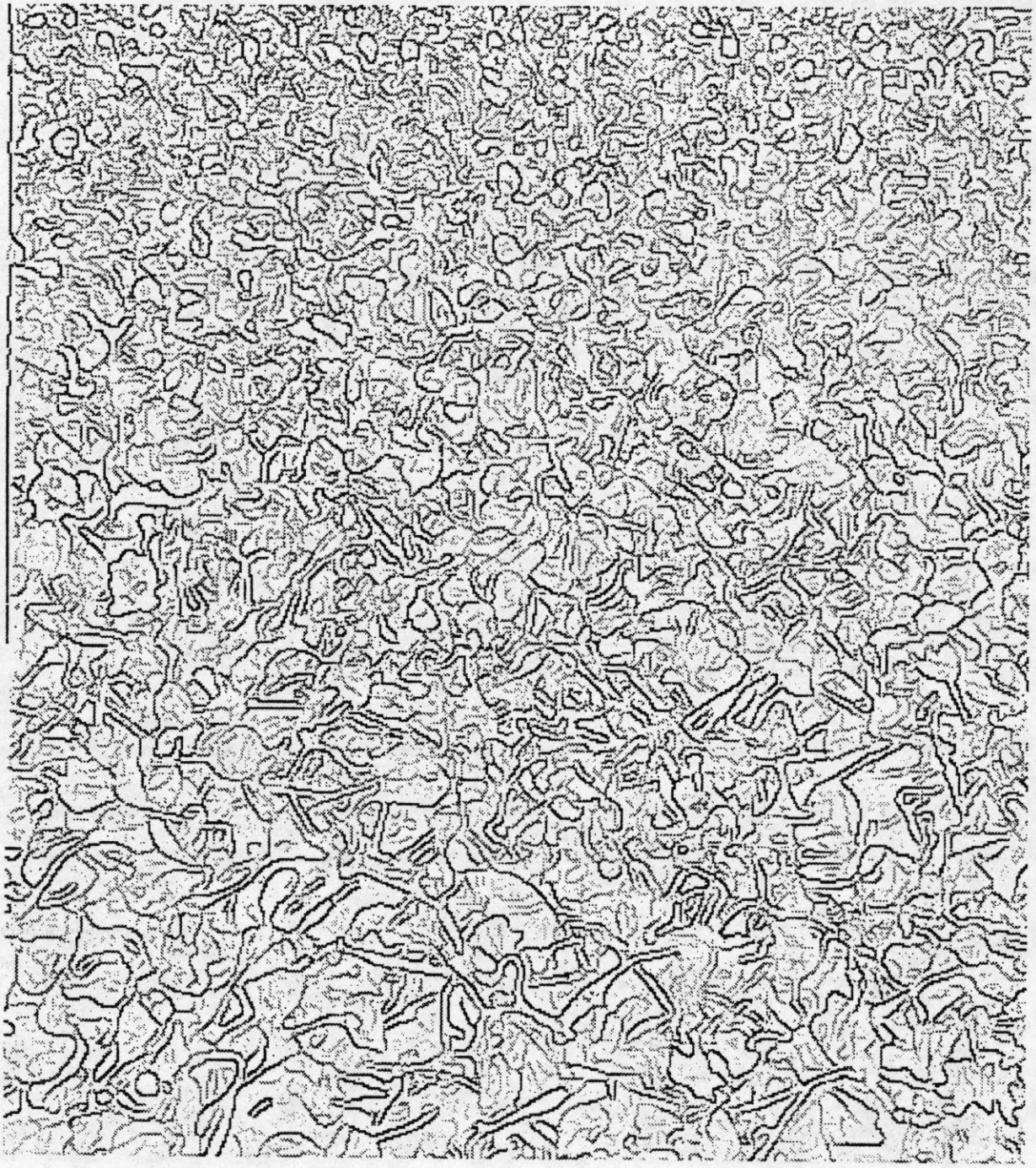
Figure 3

Edges extracted from several texture images. Only a subset of the detected edges are boundaries of texture elements. If edge density is to be effective in capturing the texture gradient, all edges that do not correspond to texel boundaries must be removed. Such edge removal cannot be accomplished without, in effect, performing an identification of texture elements. (a) Edges from the rock-pile image shown in Figure 5(a).



(b)

Figure 3, continued
(b) Edges from the image of birds flying over water, shown in Figure 9(a).



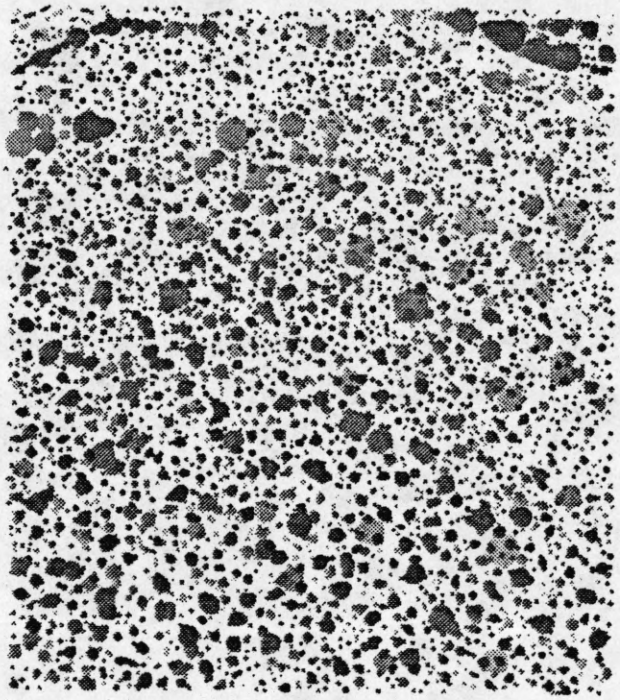
(c)

Figure 3, continued
(c) Edges from the water hyacinths image shown in Figure 31(a).

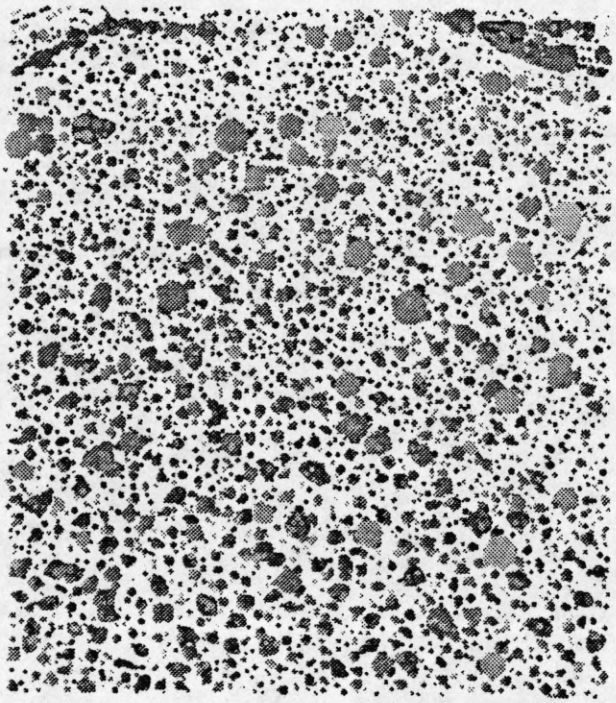
(d)

Figure 3, continued
(d) Edges from the river image shown in Figure 33(a).

(a)



(b)

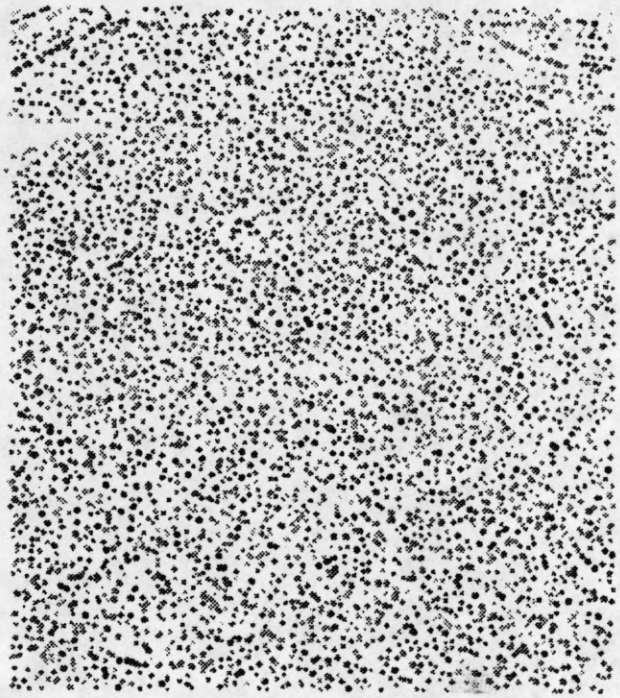


(c)

Figure 4

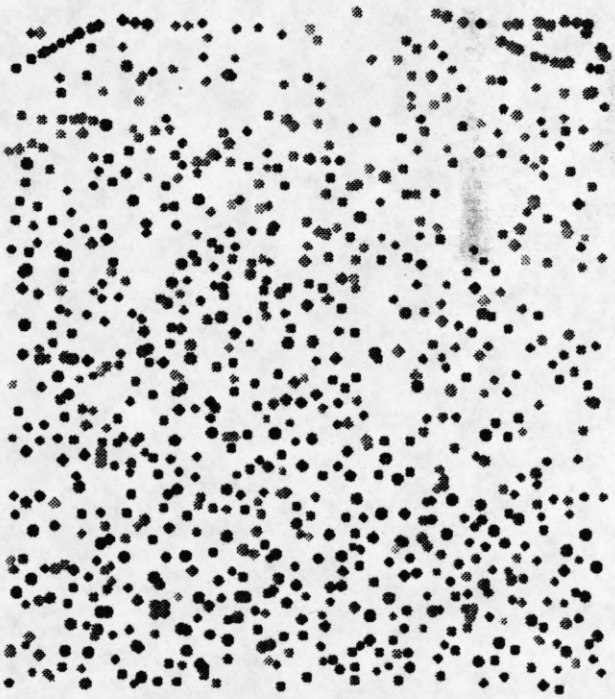
Details of the disk-fitting process for a rock-pile image. (a) The rock pile. (b) and (c) Disks corresponding to positive-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. At pixel locations covered by several disks, (b) displays the disk of higher contrast and (c) displays the disk of lower contrast. Following pages show the disks detected at each $\nabla^2 G$ filter size. The set of disks shown in (b) and (c) includes all of the disks from (e), (g), (i) and (k).

(d)



(e)

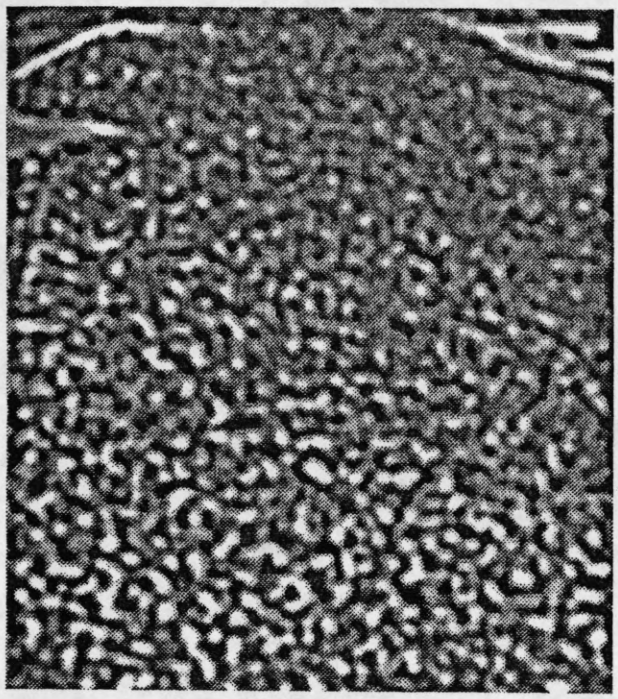
(f)



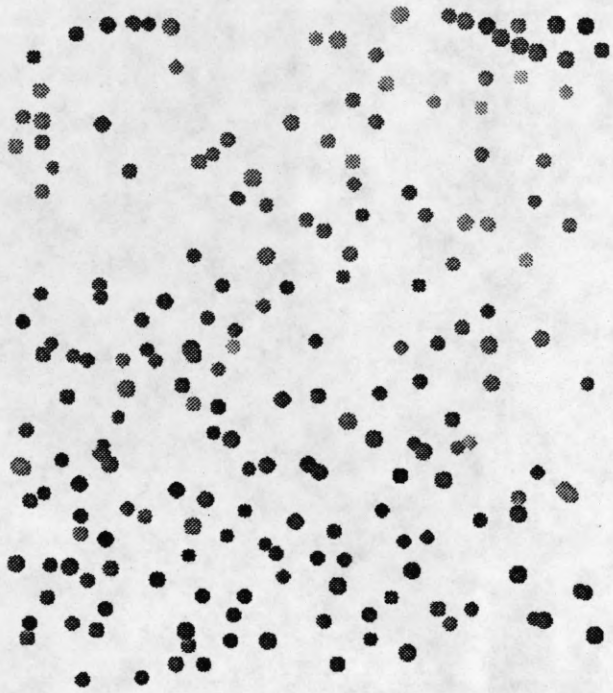
(g)

Figure 4, continued

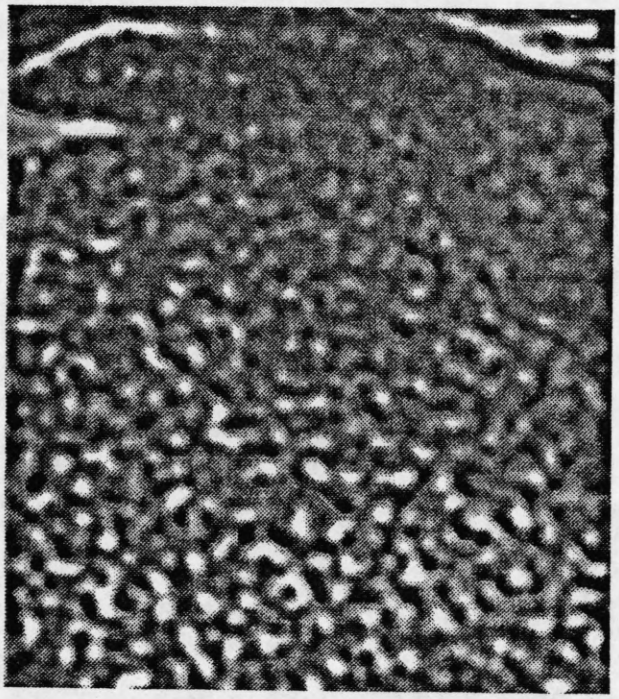
(d) Convolution of the rock-pile image with a $\nabla^2 G$ filter of size $\sigma=\sqrt{2}$; the center lobe of the $\nabla^2 G$ filter has a diameter of 4 pixels. (e) Disks detected at this filter size. The disk diameters range from 2 to 6 pixels. (f) Convolution of the rock-pile image with a $\nabla^2 G$ filter of size $\sigma=2\sqrt{2}$; the center lobe of the $\nabla^2 G$ filter has a diameter of 8 pixels. (g) Disks detected at this filter size. The disk diameters range from 6 to 10 pixels.



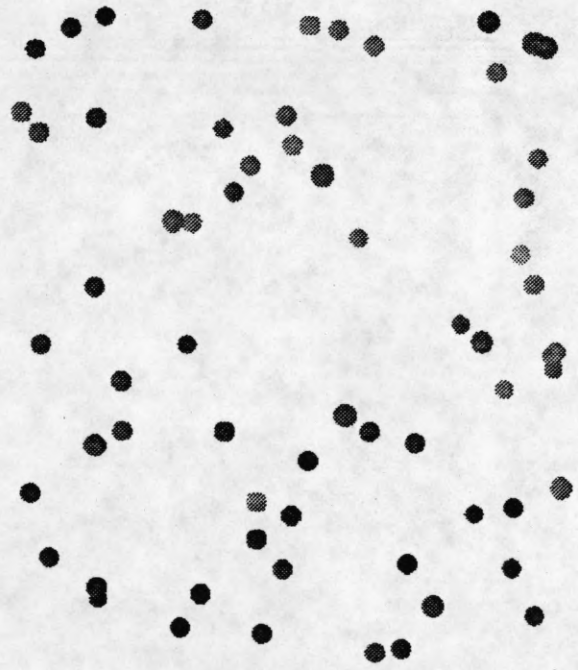
(h)



(i)



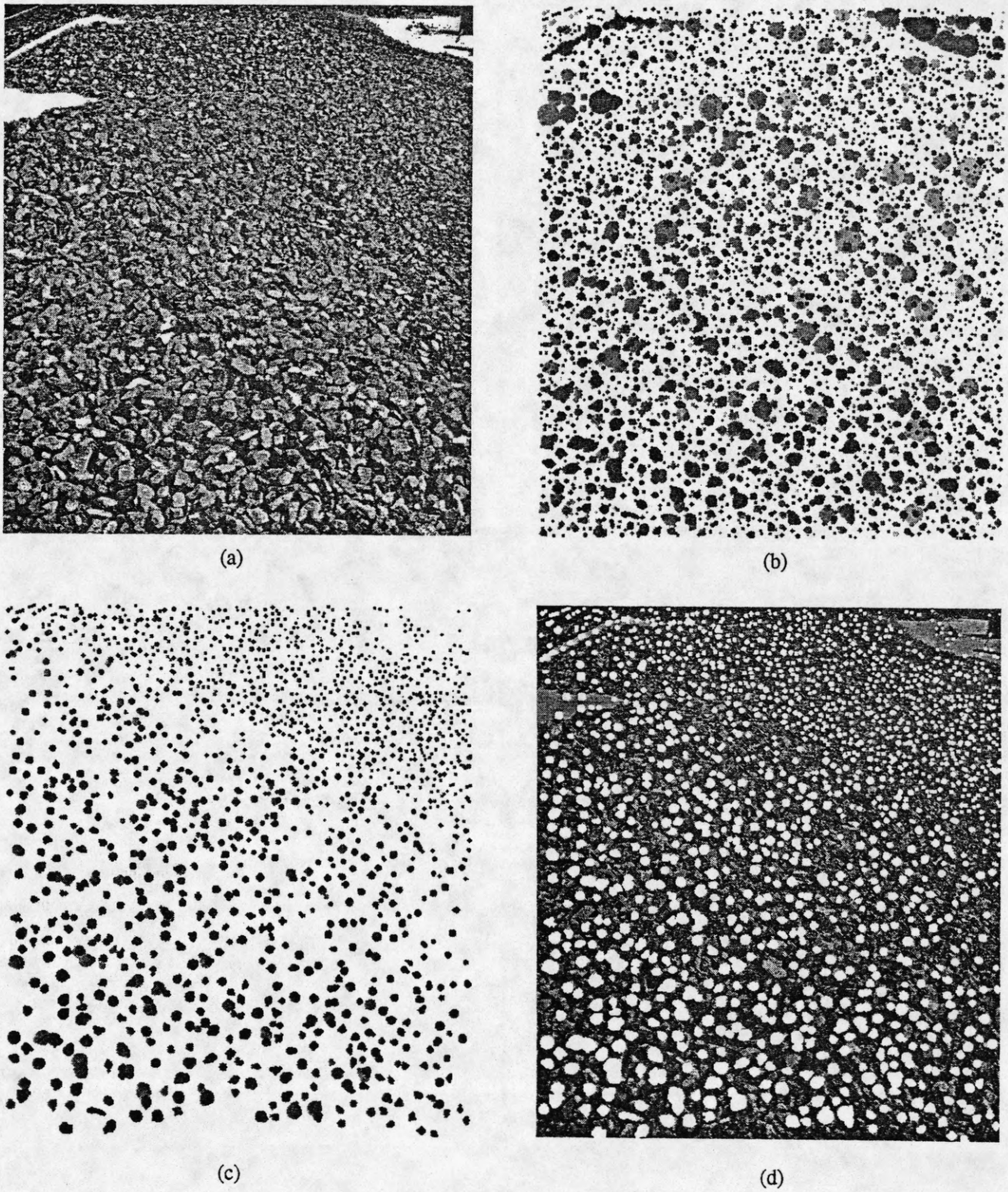
(j)



(k)

Figure 4, continued

(h) Convolution of the rock-pile image with a $\nabla^2 G$ filter of size $\sigma=3\sqrt{2}$; the center lobe of the $\nabla^2 G$ filter has a diameter of 12 pixels. (i) Disks detected at this filter size. The disk diameters range from 10 to 14 pixels. (j) Convolution of the rock-pile image with a $\nabla^2 G$ filter of size $\sigma=4\sqrt{2}$; the center lobe of the $\nabla^2 G$ filter has a diameter of 16 pixels. (k) Disks detected at this filter size. The disk diameters range from 14 to 18 pixels.



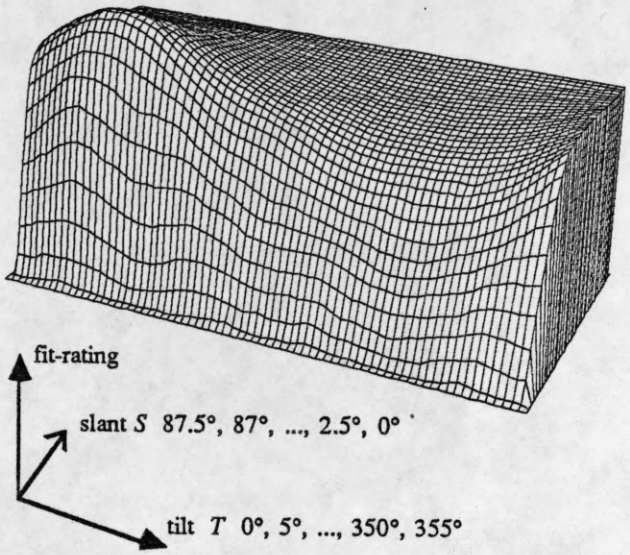
(a) (b)

(c) (d)

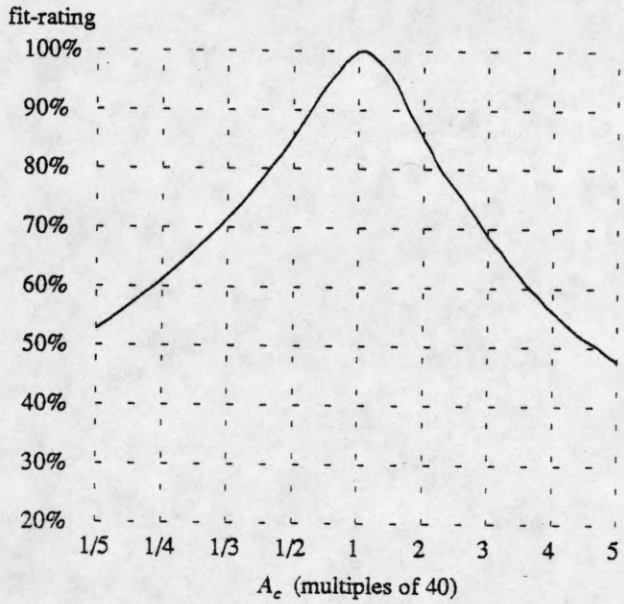
Figure 5

(a) A rock pile. (b) Disks corresponding to positive-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. (c) Extracted texels. These are all regions (sets of overlapping disks) having area within a factor of two of the area expected by the best planar fit (A_c 40, slant 62.5°, tilt 65°). The texels that fit the plane most closely are printed darkest. (d) The texels superimposed on a dark reproduction of the original.

(e)



(f)



(g)

Figure 5, continued (A rock pile; positive-contrast regions)
(e) Synthetic image to illustrate the planar fit A_c 40, slant 62.5°, tilt 65°. (f) and (g) Ratings of various possible planar fits. In (f) slant and tilt are varied while A_c is constant at 40. In (g) A_c is varied while slant and tilt are constant at 62.5° and 65° respectively.

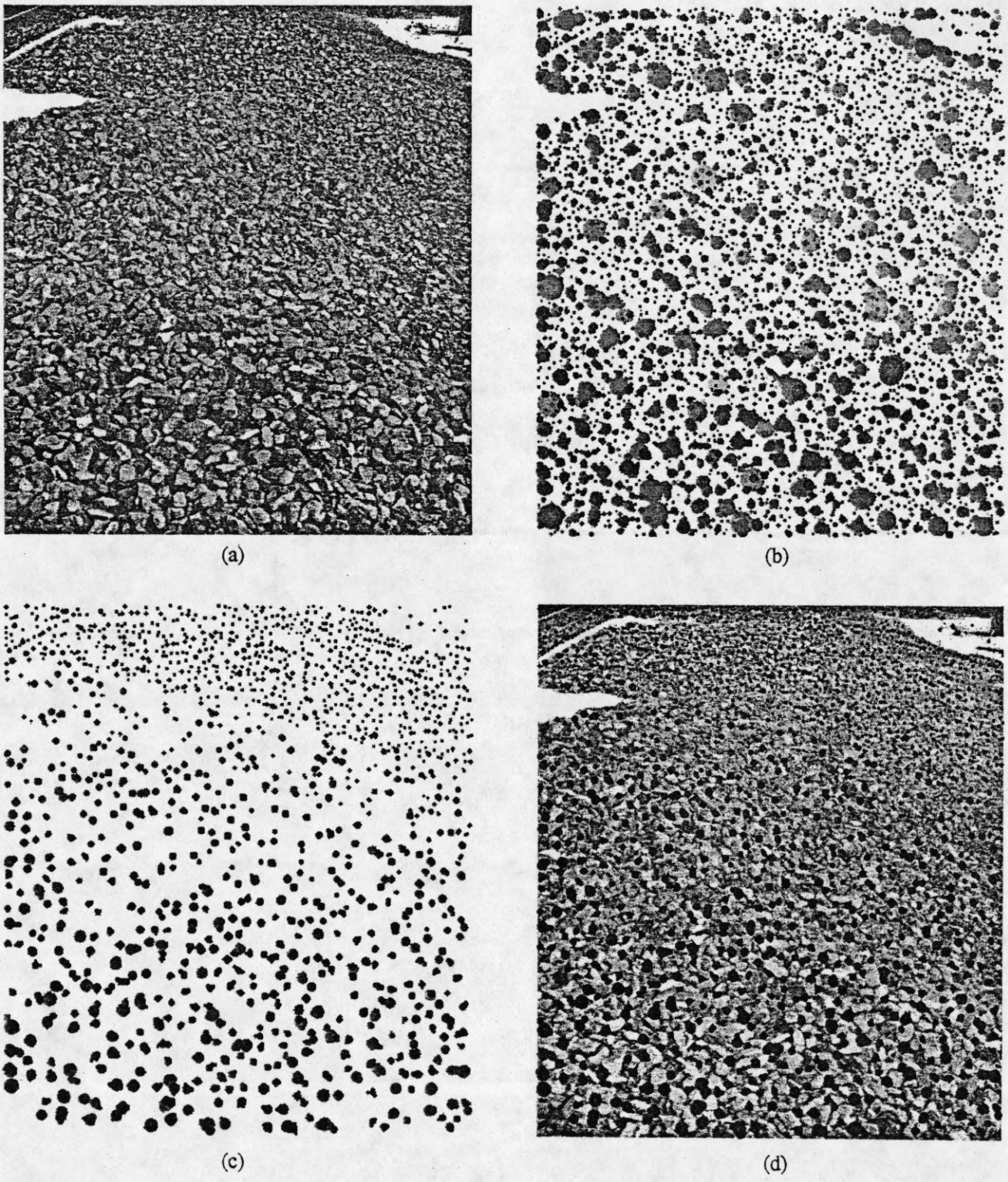
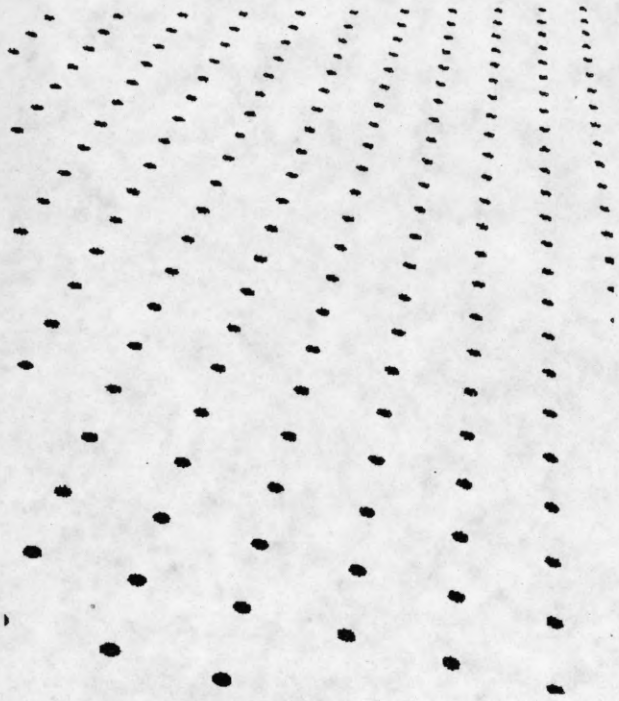
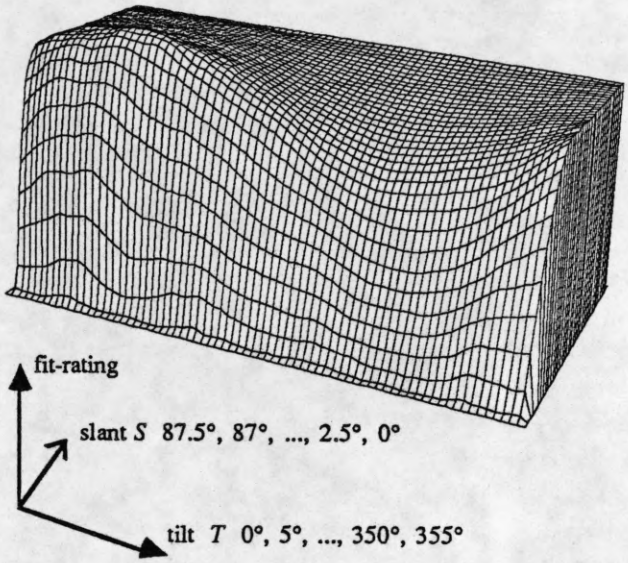


Figure 6

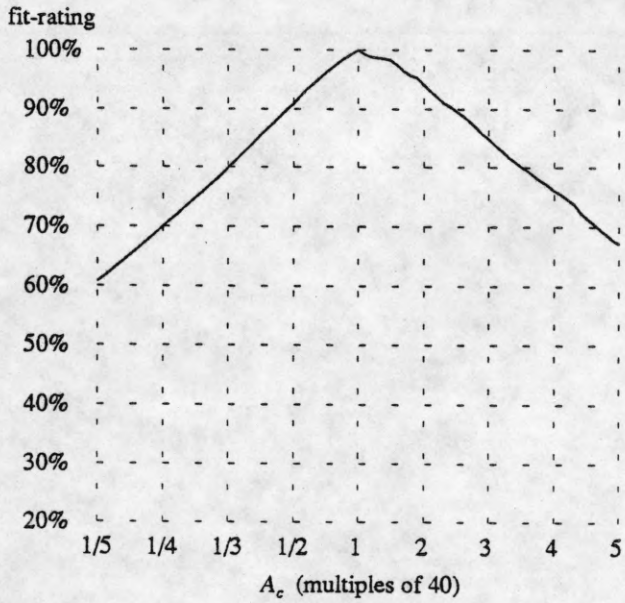
(a) A rock pile. (b) Disks corresponding to negative-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. (c) Extracted texels. These are all regions (sets of overlapping disks) having area within a factor of two of the area expected by the best planar fit (A_c 40, slant 60°, tilt 75°). The texels that fit the plane most closely are printed darkest. (d) The texels superimposed on a bright reproduction of the original.



(e)



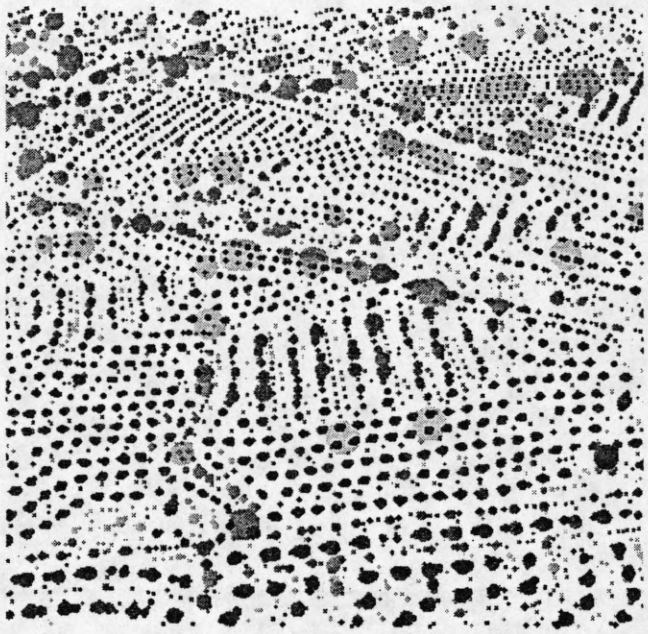
(f)



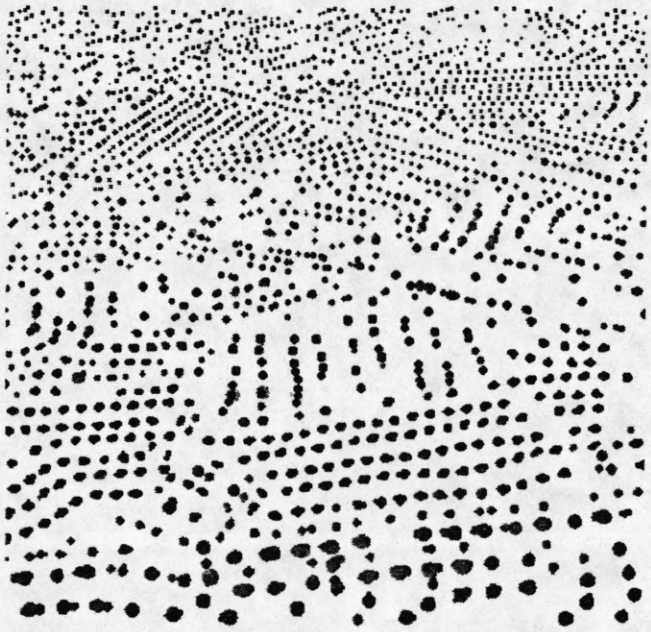
(g)

Figure 6, continued (A rock pile; negative-contrast regions)
(e) Synthetic image to illustrate the planar fit A_c 40, slant 60°, tilt 75°. (f) and (g) Ratings of various possible planar fits. In (f) slant and tilt are varied while A_c is constant at 40. In (g) A_c is varied while slant and tilt are constant at 60° and 75° respectively.

(a)



(b)

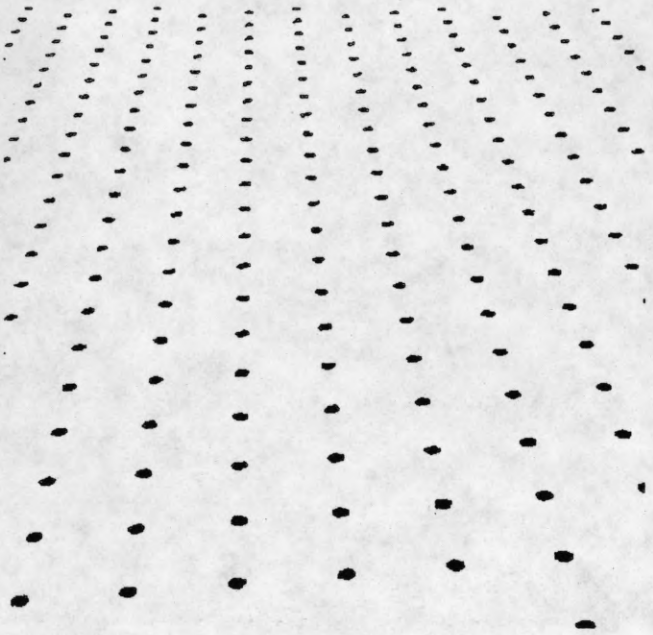


(c)

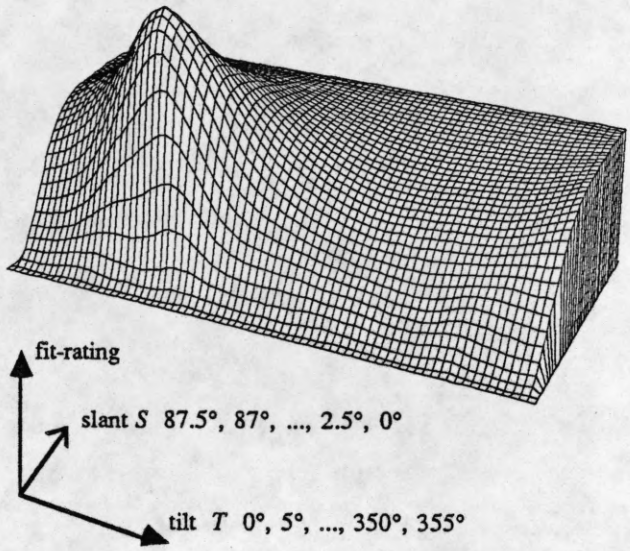
(d)

Figure 7

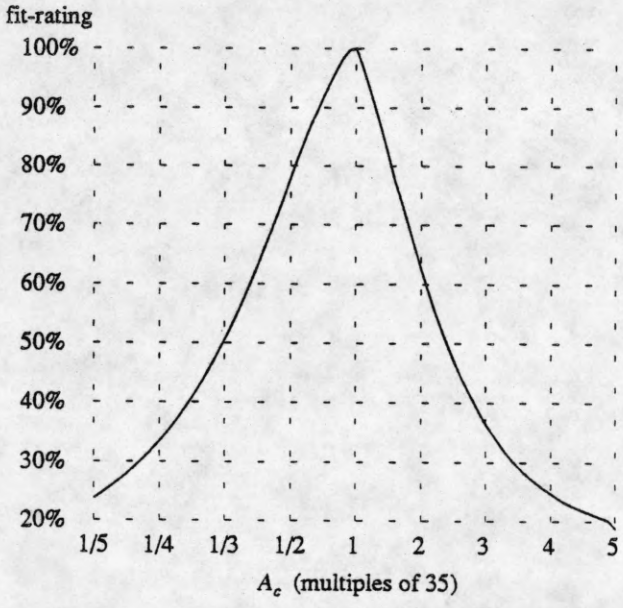
(a) Aerial view of Littown, Pennsylvania. (b) Disks corresponding to positive-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. (c) Extracted texels. These are all regions (sets of overlapping disks) having area within a factor of two of the area expected by the best planar fit (A_c 35, slant 62.5°, tilt 95°). The texels that fit the plane most closely are printed darkest. (d) The texels superimposed on a dark reproduction of the original.



(e)



(f)

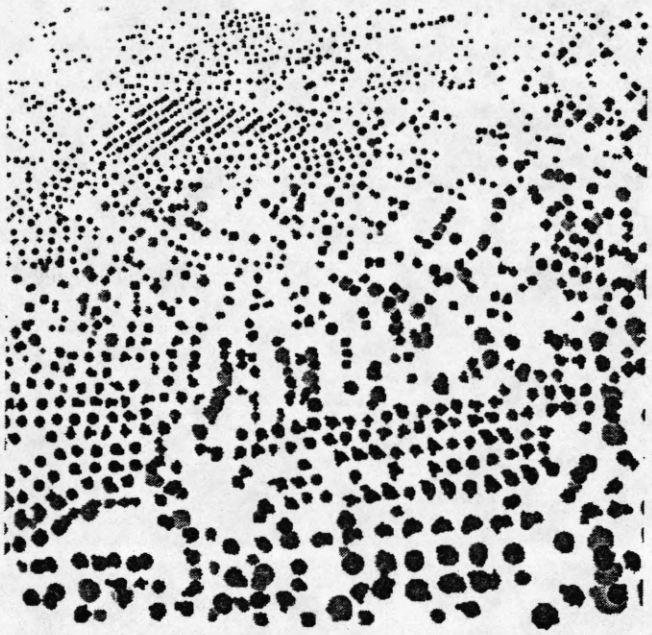


(g)

Figure 7, continued (Aerial view of Littown, Pennsylvania; positive-contrast regions)
(e) Synthetic image to illustrate the planar fit A_c 35, slant 62.5°, tilt 95°. (f) and (g) Ratings of various possible planar fits. In (f) slant and tilt are varied while A_c is constant at 35. In (g) A_c is varied while slant and tilt are constant at 62.5° and 95° respectively.

(a)

(b)

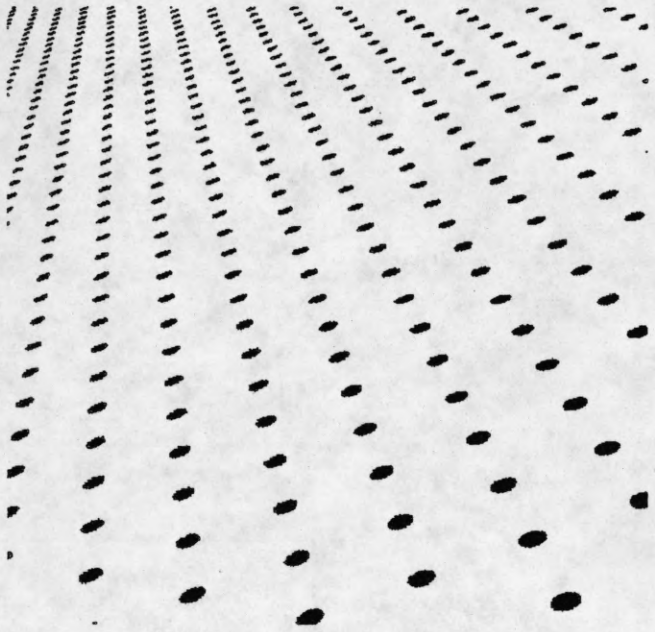


(c)

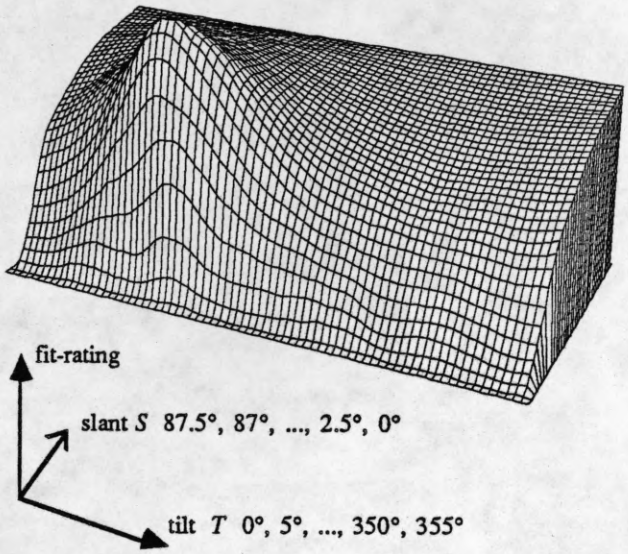
(d)

Figure 8

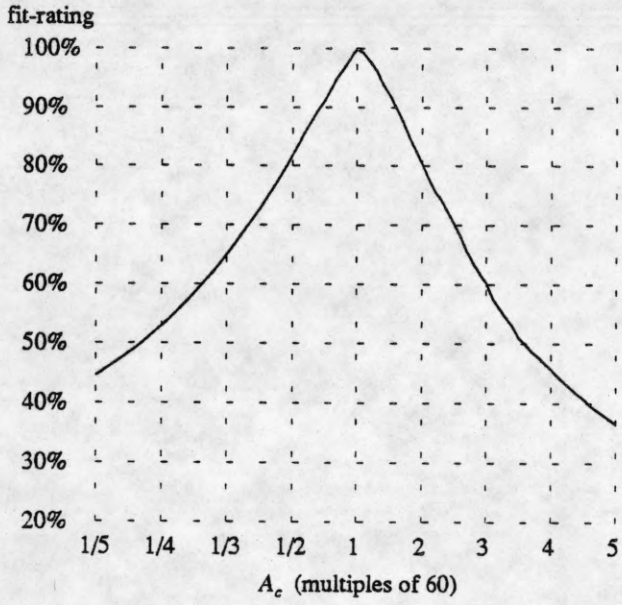
(a) Aerial view of Littown, Pennsylvania. (b) Disks corresponding to negative-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. (c) Extracted texels. These are all regions (sets of overlapping disks) having area within a factor of two of the area expected by the best planar fit (A_c 60, slant 67.5°, tilt 110°). The texels that fit the plane most closely are printed darkest. (d) The texels superimposed on a bright reproduction of the original.



(e)

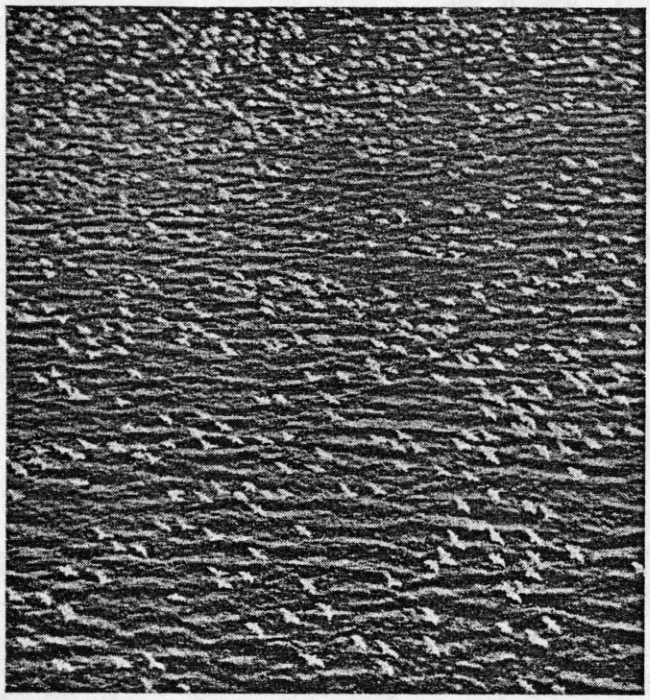


(f)

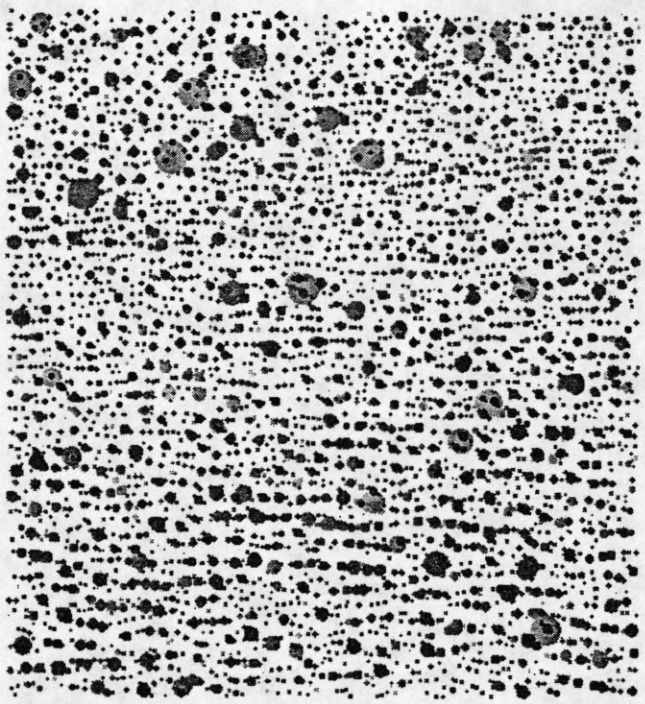


(g)

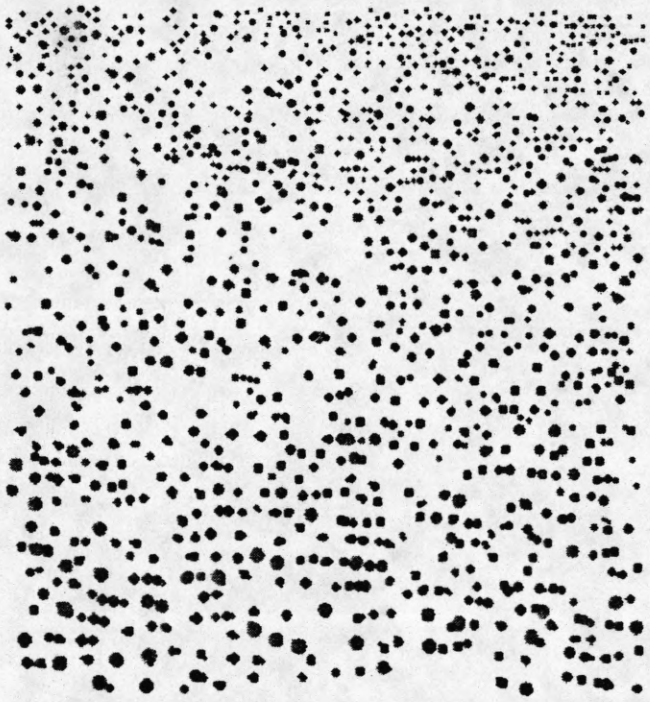
Figure 8, continued (Aerial view of Littown, Pennsylvania; negative-contrast regions)
(e) Synthetic image to illustrate the planar fit A_c 60, slant 67.5°, tilt 110°. (f) and (g) Ratings of various possible planar fits. In (f) slant and tilt are varied while A_c is constant at 60. In (g) A_c is varied while slant and tilt are constant at 67.5° and 110° respectively.



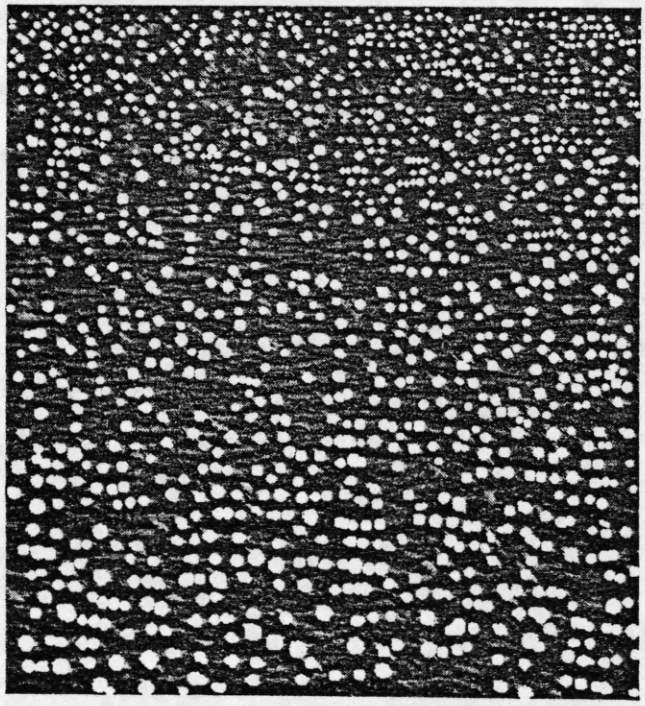
(a)



(b)



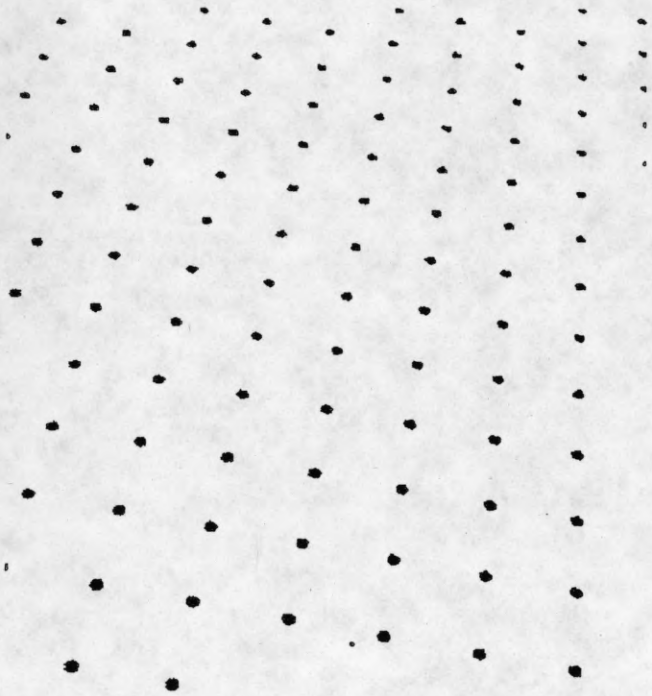
(c)



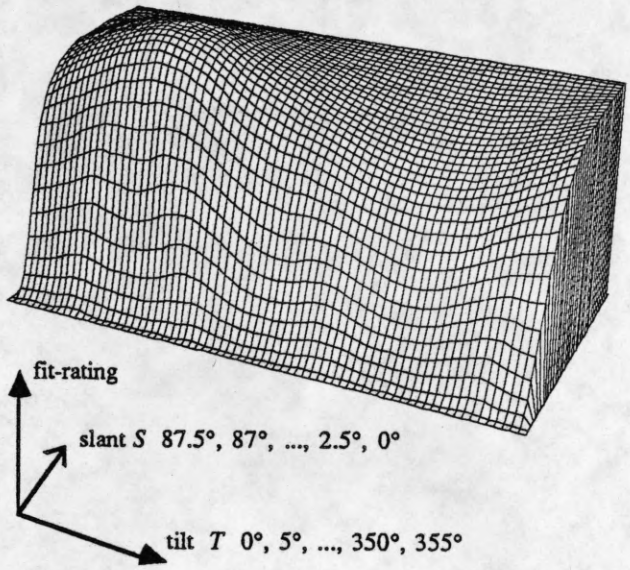
(d)

Figure 9

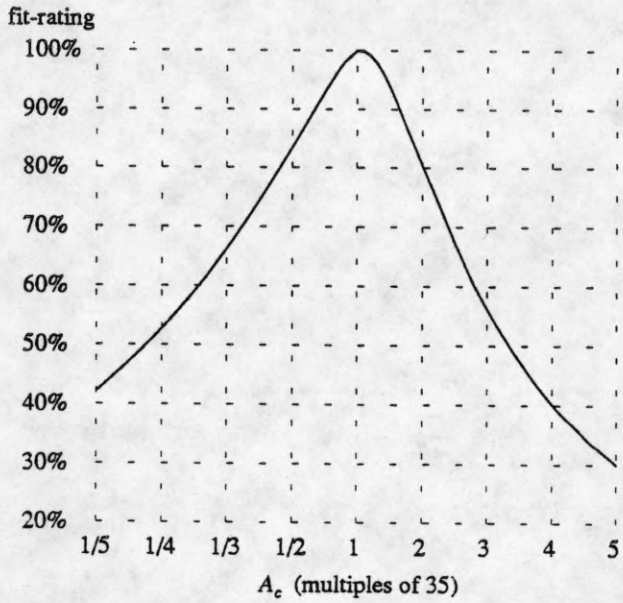
(a) Snow geese over Back Bay, Virginia. (b) Disks corresponding to positive-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. (c) Extracted texels. These are all regions (sets of overlapping disks) having area within a factor of two of the area expected by the best planar fit (A_c 35, slant 45°, tilt 80°). The texels that fit the plane most closely are printed darkest. (d) The texels superimposed on a dark reproduction of the original.



(e)



(f)



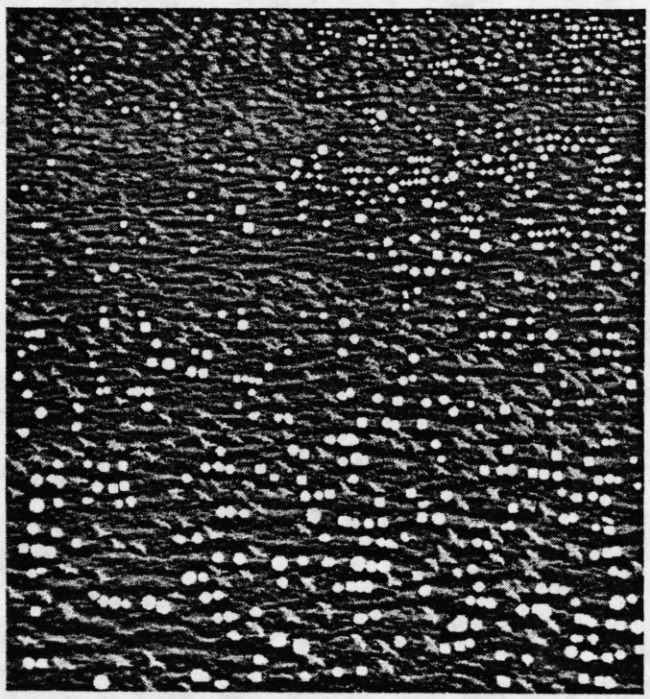
(g)

Figure 9, continued (Snow geese over Back Bay, Virginia; positive-contrast regions)
(e) Synthetic image to illustrate the planar fit A_c 35, slant 45°, tilt 80°. (f) and (g) Ratings of various possible planar fits. In (f) slant and tilt are varied while A_c is constant at 35. In (g) A_c is varied while slant and tilt are constant at 45° and 80° respectively.

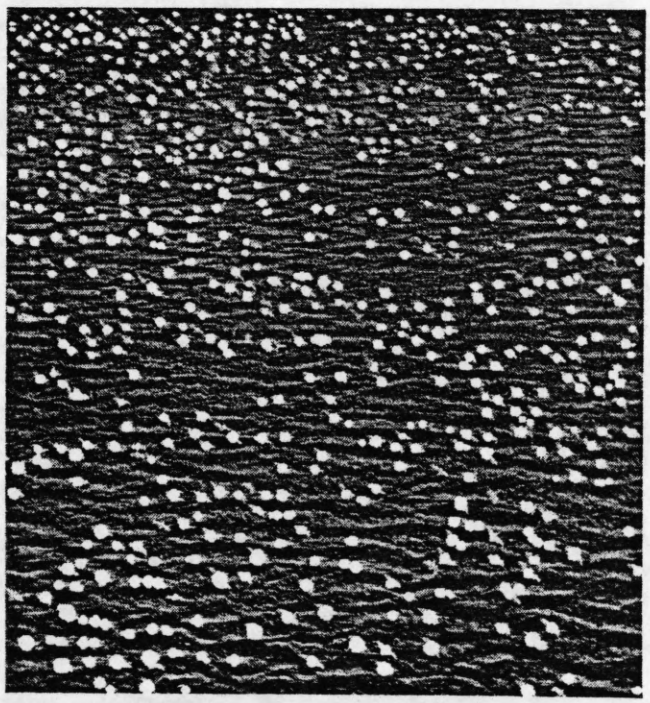
```
Regions within -2 to 2 of center-area 35, slant 45, tilt 80

Histogram of mean_grays of regions in ../birds.regions.fit.2:
320-329      2 *
330-339      5 **
340-349     28 ************
350-359     51 **********************
360-369     90 ****************************************
370-379    102 *********************************************
380-389    120 ****************************************************
390-399     98 *******************************************
400-409     87 ***************************************
410-419     68 ******************************
420-429     74 *********************************
430-439     72 ********************************
440-449    103 *********************************************
450-459    105 **********************************************
460-469    144 ****************************************************************
470-479    123 ******************************************************
480-489     91 ****************************************
490-499     64 ****************************
500-509     21 *********
510-519      5 **
Total number of entries: 1453;  average: 426.9;  maximum: 518.1
```

(h)

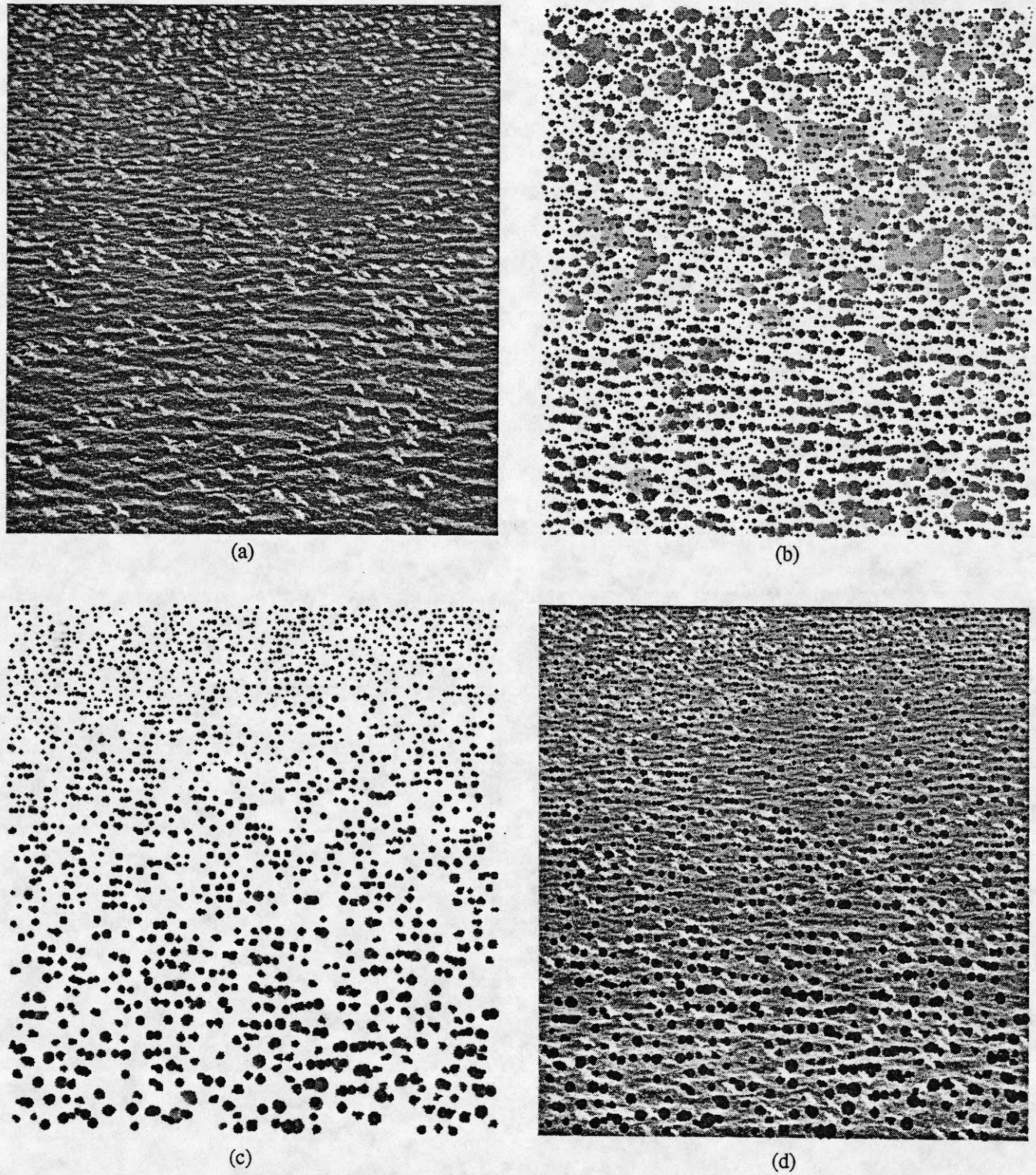


(i)



(j)

Figure 9, continued (Snow geese over Back Bay, Virginia; positive-contrast regions) Categorizing the texels into a field of birds and a field of wave crests. (h) Histogram of the average gray-level of the texels from (c). (i) Wave crests: texels with average gray-level less than 415. (j) Birds: texels with average gray-level greater than 415. A better region detector would reduce the need for gray-level based categorization of texels: with more accurate detection of elongated regions, the birds and waves could be recognized as separate texture fields based on shape properties of the texels.

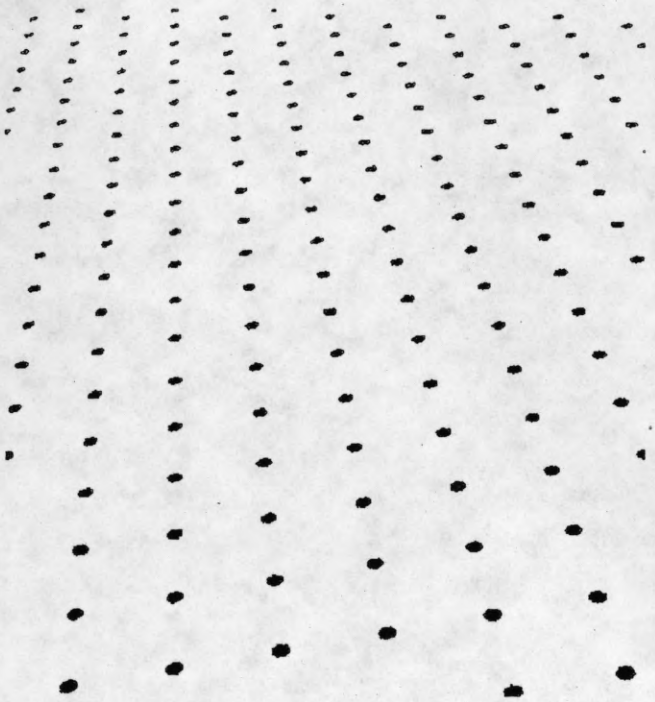


(a) (b)

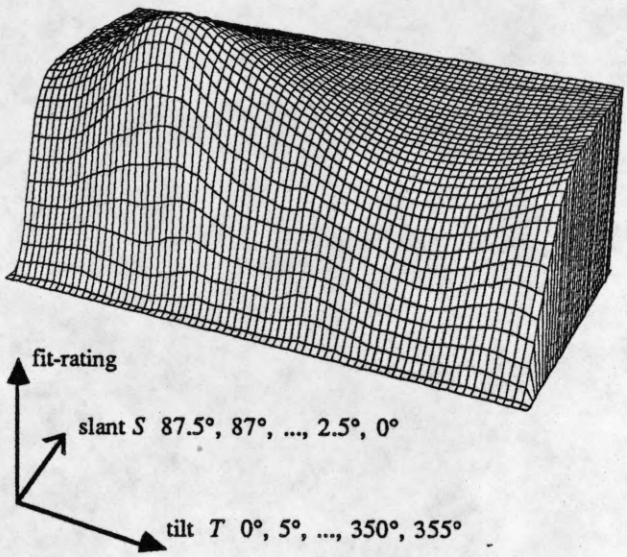
(c) (d)

Figure 10

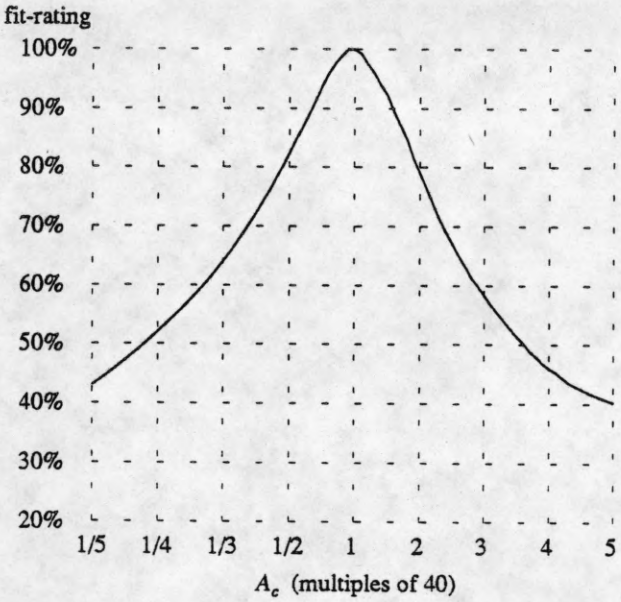
(a) Snow geese over Back Bay, Virginia. (b) Disks corresponding to negative-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. (c) Extracted texels. These are all regions (sets of overlapping disks) having area within a factor of two of the area expected by the best planar fit (A_c 40, slant 57.5°, tilt 100°). The texels that fit the plane most closely are printed darkest. (d) The texels superimposed on a bright reproduction of the original.



(e)



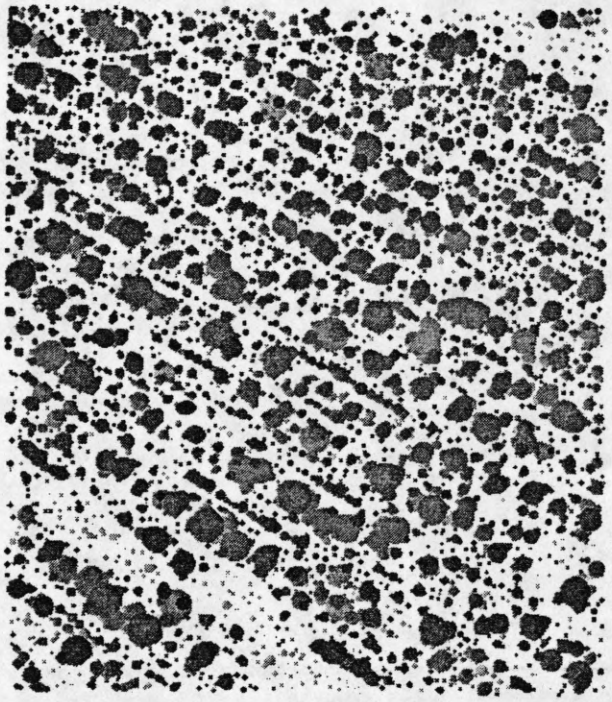
(f)



(g)

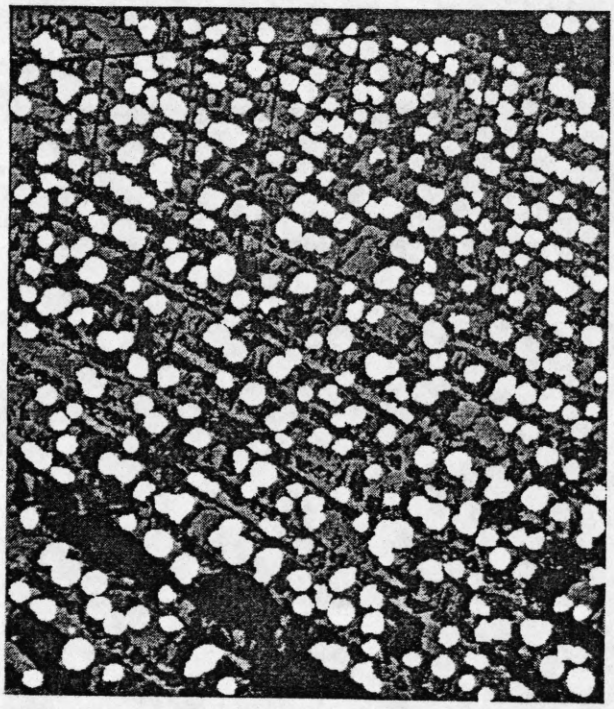
Figure 10, continued (Snow geese over Back Bay, Virginia; negative-contrast regions)
(e) Synthetic image to illustrate the planar fit A_c 40, slant 57.5°, tilt 100°. (f) and (g) Ratings of various possible planar fits. In (f) slant and tilt are varied while A_c is constant at 40. In (g) A_c is varied while slant and tilt are constant at 57.5° and 100° respectively.

(a)



(b)

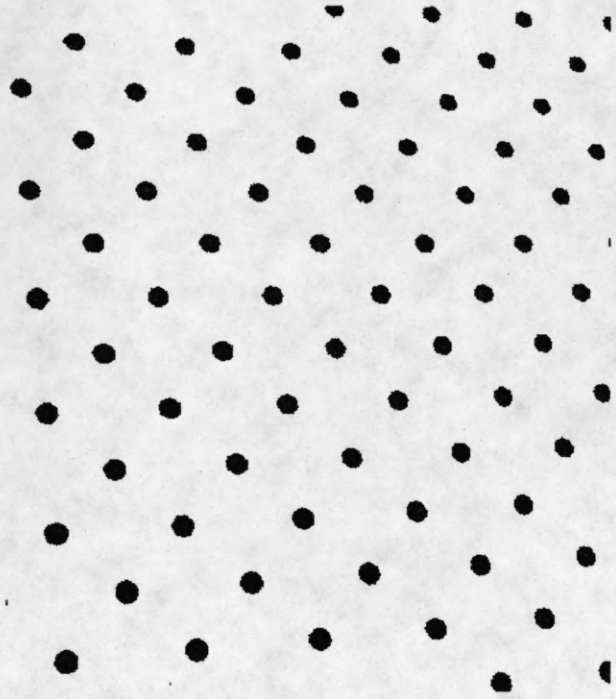
(c)



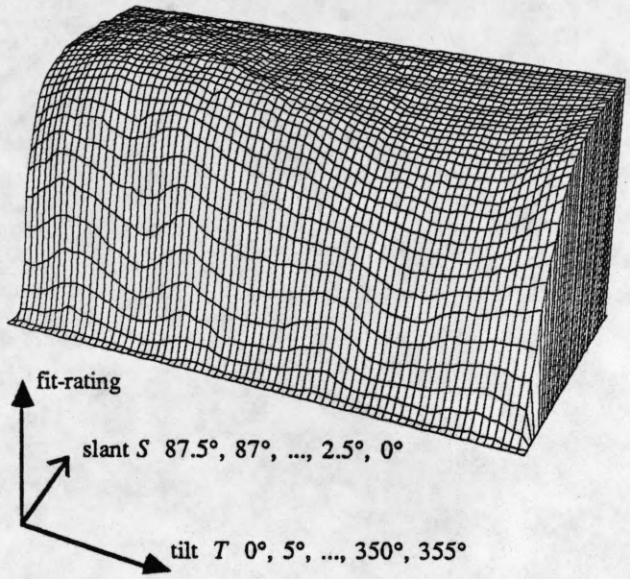
(d)

Figure 11

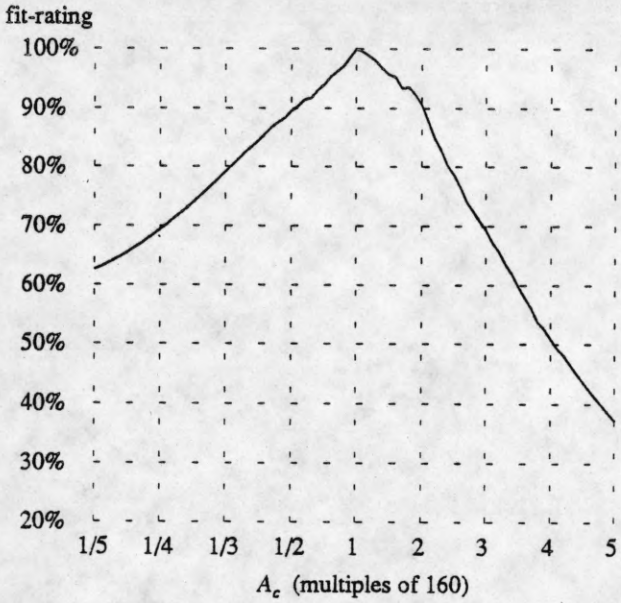
(a) Muslims at a mosque. (b) Disks corresponding to positive-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. (c) Extracted texels. These are all regions (sets of overlapping disks) having area within a factor of two of the area expected by the best planar fit (A_c 160, slant 27.5°, tilt 50°). The texels that fit the plane most closely are printed darkest. (d) The texels superimposed on a dark reproduction of the original.



(e)

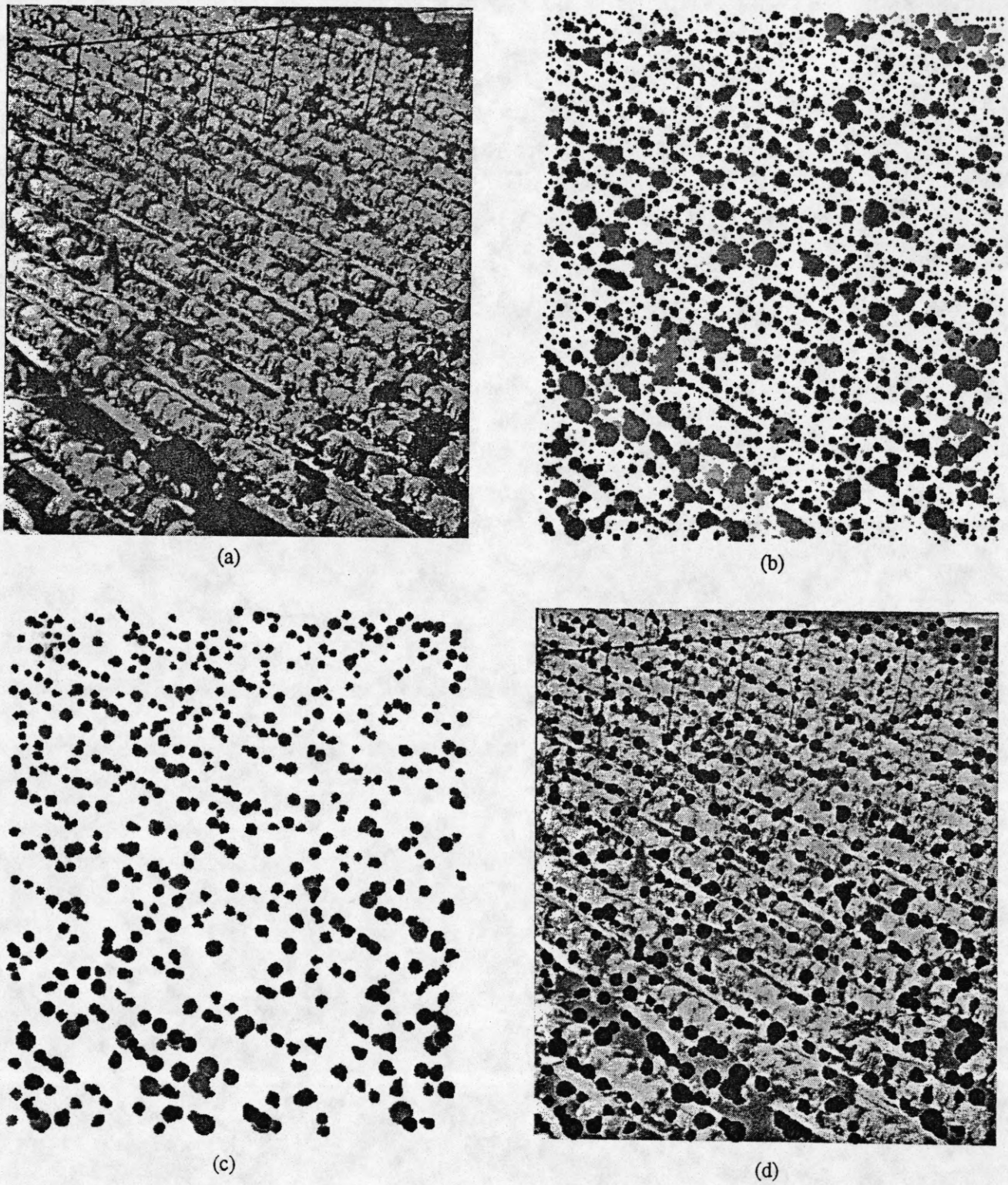


(f)



(g)

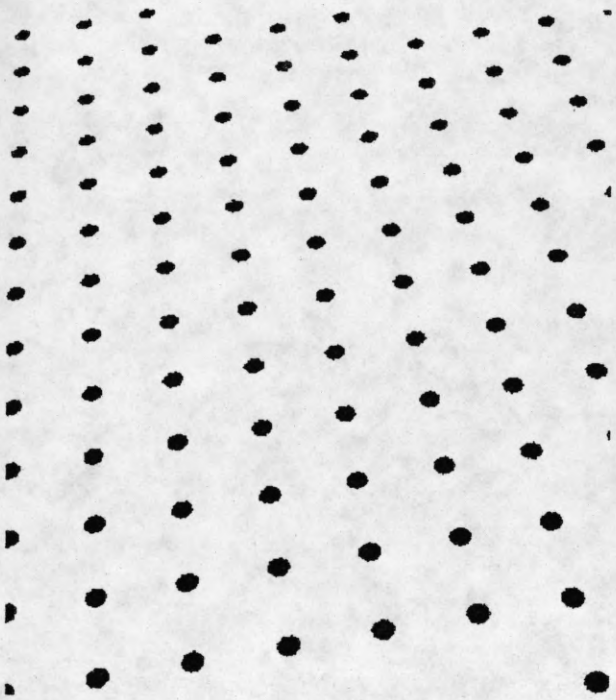
Figure 11, continued (Muslims at a mosque; positive-contrast regions)
(e) Synthetic image to illustrate the planar fit A_c 160, slant 27.5°, tilt 50°. (f) and (g) Ratings of various possible planar fits. In (f) slant and tilt are varied while A_c is constant at 160. In (g) A_c is varied while slant and tilt are constant at 27.5° and 50° respectively.



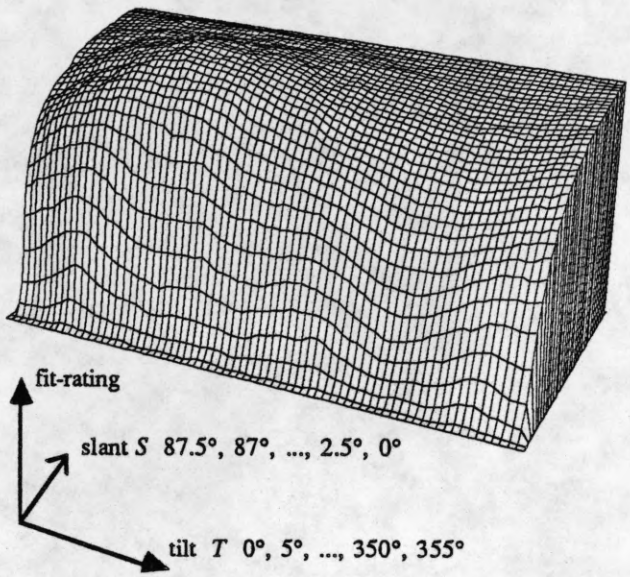
(a) (b) (c) (d)

Figure 12

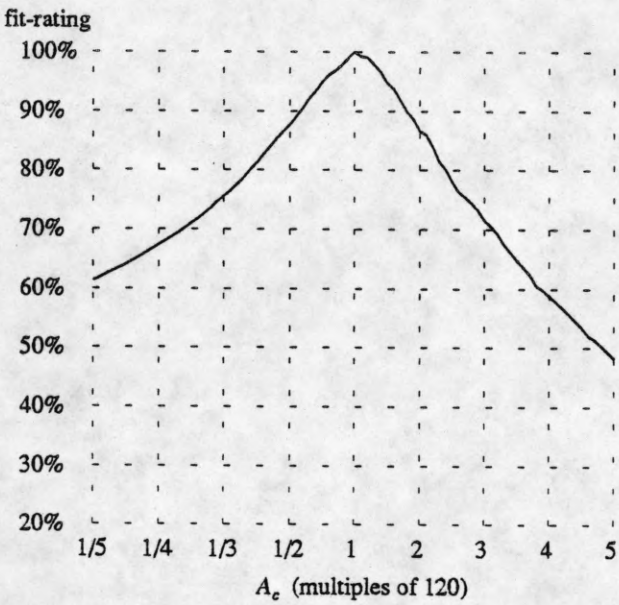
(a) Muslims at a mosque. (b) Disks corresponding to negative-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. (c) Extracted texels. These are all regions (sets of overlapping disks) having area within a factor of two of the area expected by the best planar fit (A_c 120, slant 42.5°, tilt 100°). The texels that fit the plane most closely are printed darkest. (d) The texels superimposed on a bright reproduction of the original.



(e)



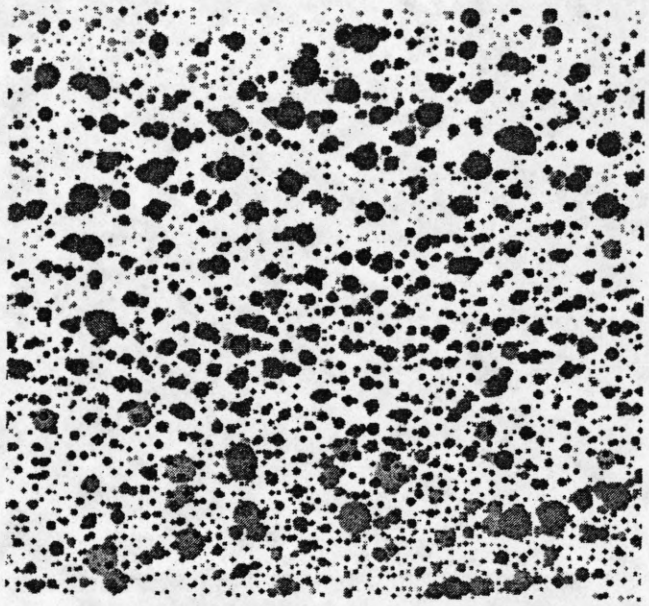
(f)



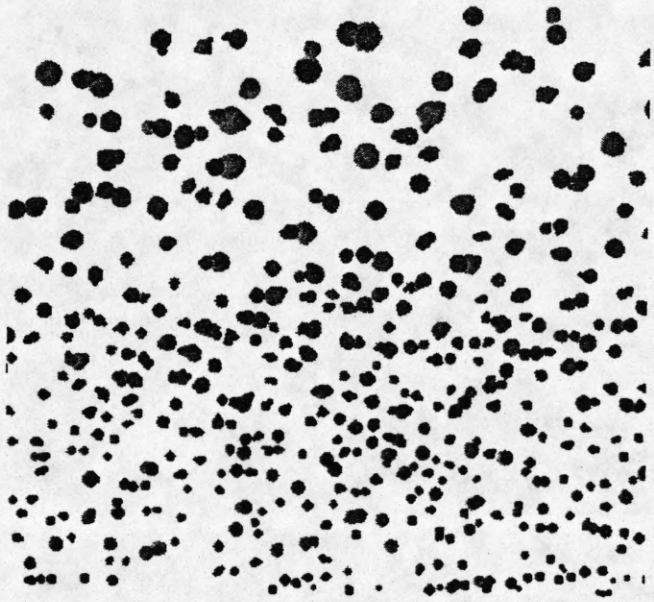
(g)

Figure 12, continued (Muslims at a mosque; negative-contrast regions)
(e) Synthetic image to illustrate the planar fit A_c 120, slant 42.5°, tilt 100°. (f) and (g) Ratings of various possible planar fits. In (f) slant and tilt are varied while A_c is constant at 120. In (g) A_c is varied while slant and tilt are constant at 42.5° and 100° respectively.

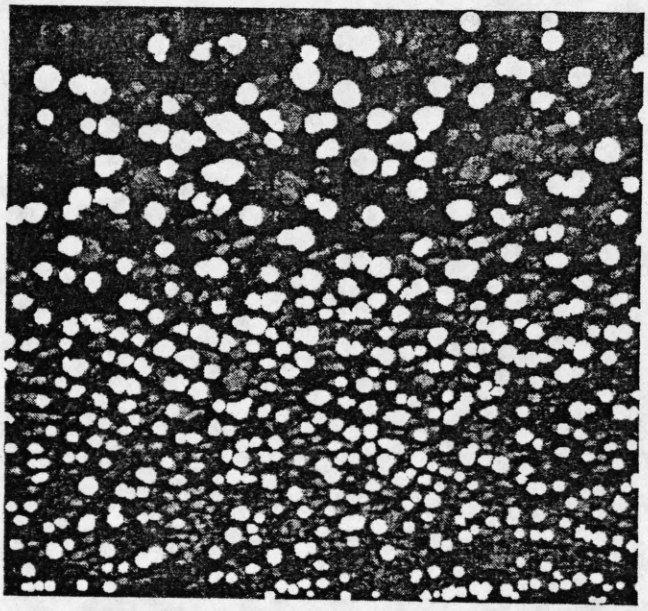
(a)



(b)



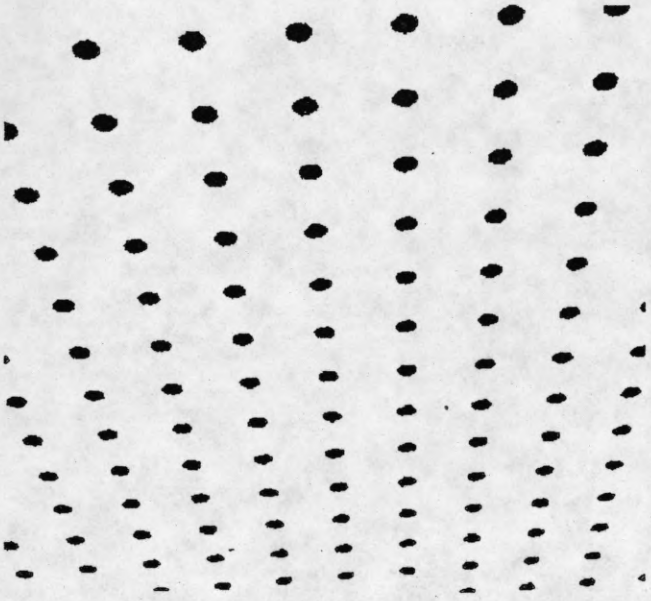
(c)



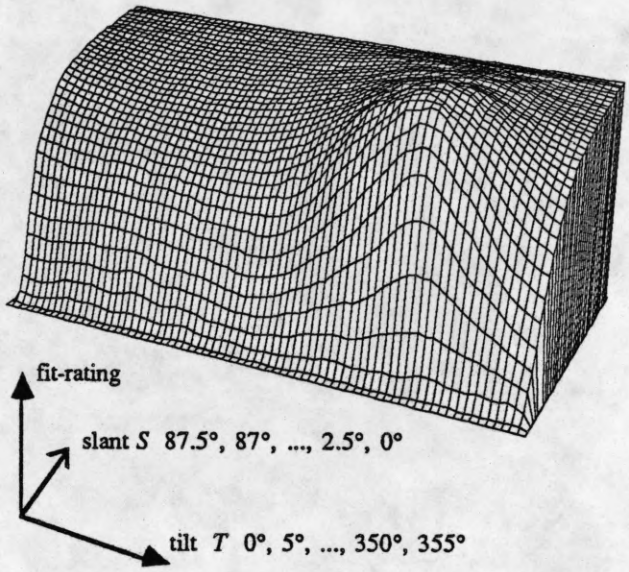
(d)

Figure 13

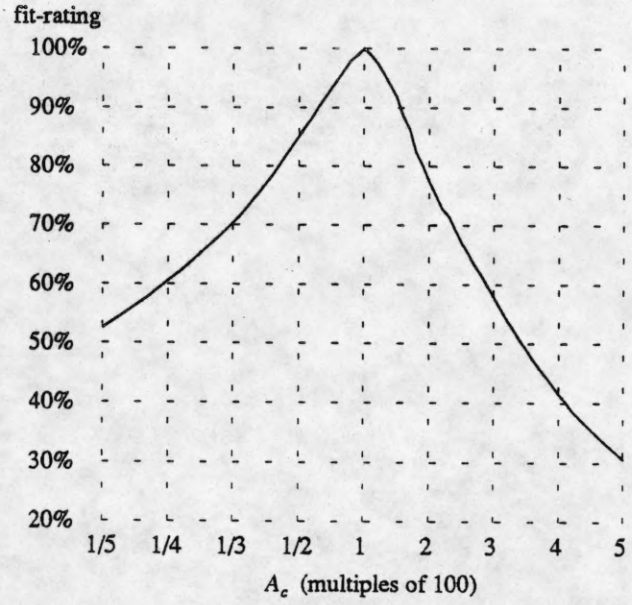
(a) Fleecy clouds. (b) Disks corresponding to positive-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. (c) Extracted texels. These are all regions (sets of overlapping disks) having area within a factor of two of the area expected by the best planar fit (A_c 100, slant 55°, tilt 275°). The texels that fit the plane most closely are printed darkest. (d) The texels superimposed on a dark reproduction of the original.



(e)



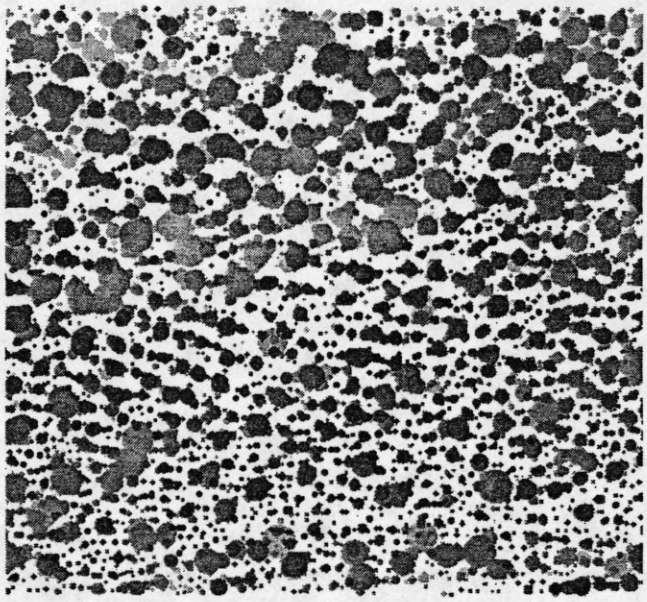
(f)



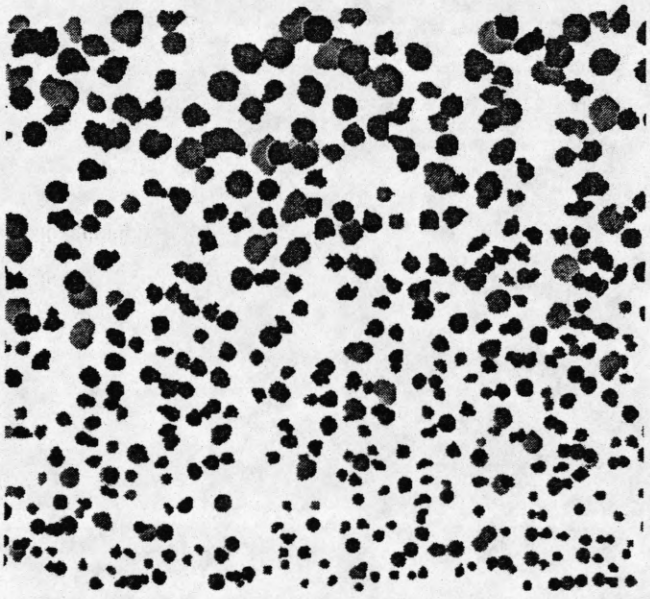
(g)

Figure 13, continued (Fleecy clouds; positive-contrast regions)
(e) Synthetic image to illustrate the planar fit A_c 100, slant 55°, tilt 275°. (f) and (g) Ratings of various possible planar fits. In (f) slant and tilt are varied while A_c is constant at 100. In (g) A_c is varied while slant and tilt are constant at 55° and 275° respectively.

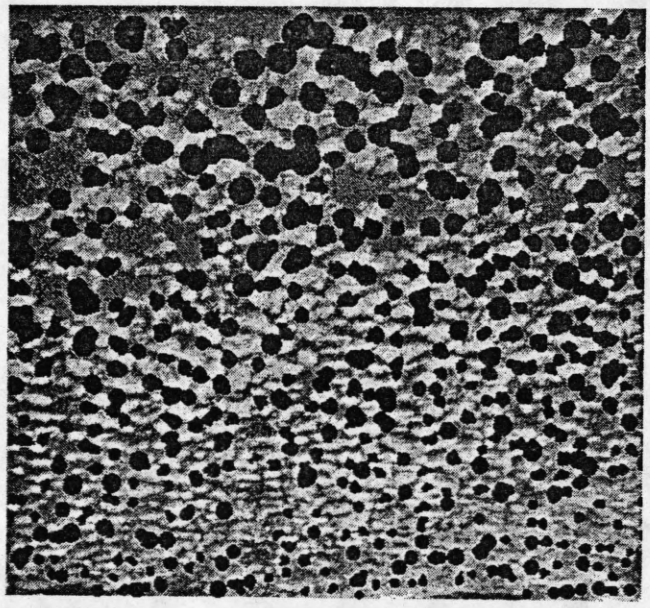
(a)



(b)



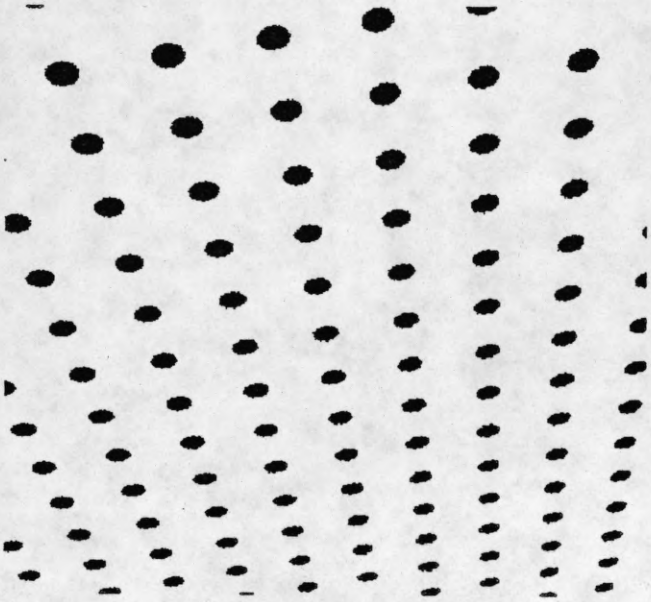
(c)



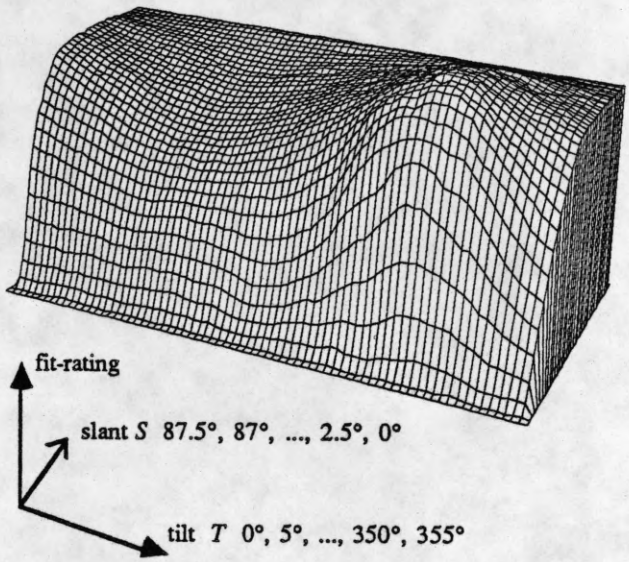
(d)

Figure 14

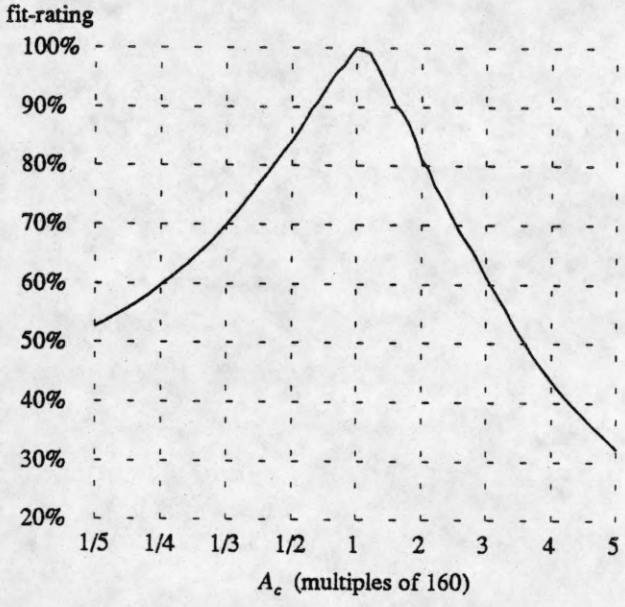
(a) Fleecy clouds. (b) Disks corresponding to negative-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. (c) Extracted texels. These are all regions (sets of overlapping disks) having area within a factor of two of the area expected by the best planar fit (A_c 160, slant 55°, tilt 280°). The texels that fit the plane most closely are printed darkest. (d) The texels superimposed on a bright reproduction of the original.



(e)

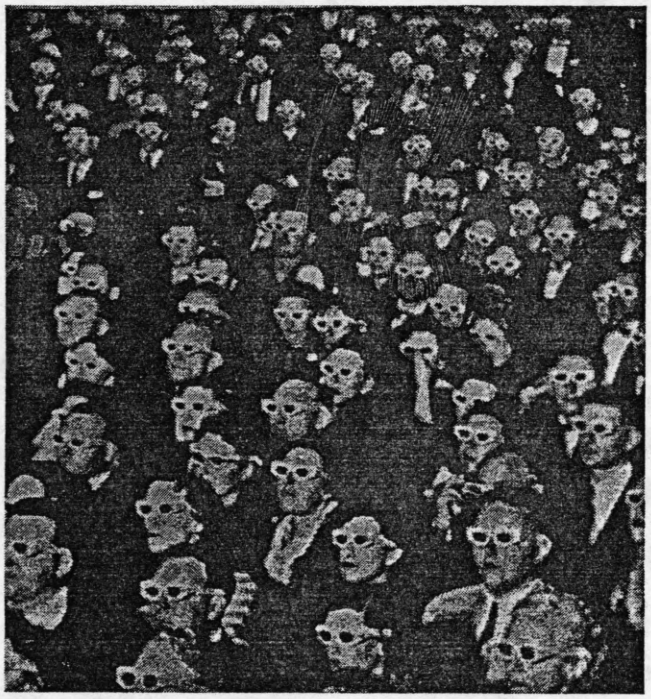


(f)

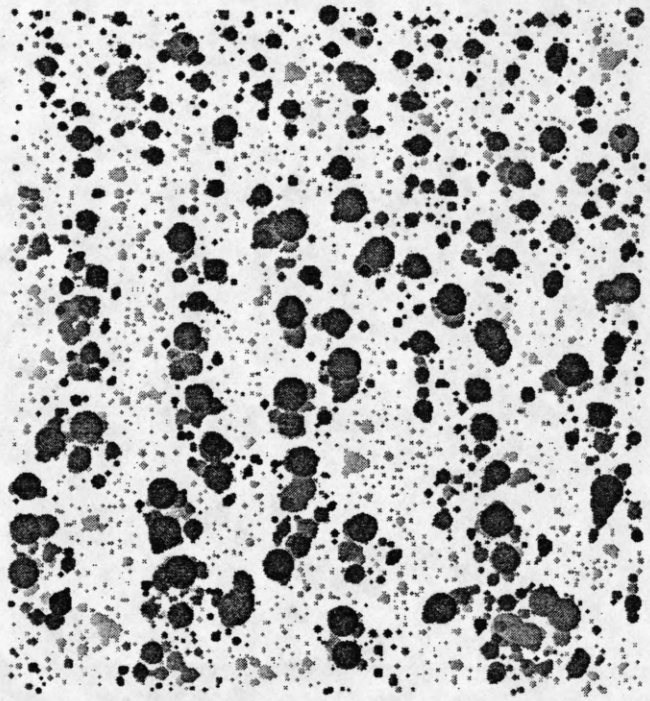


(g)

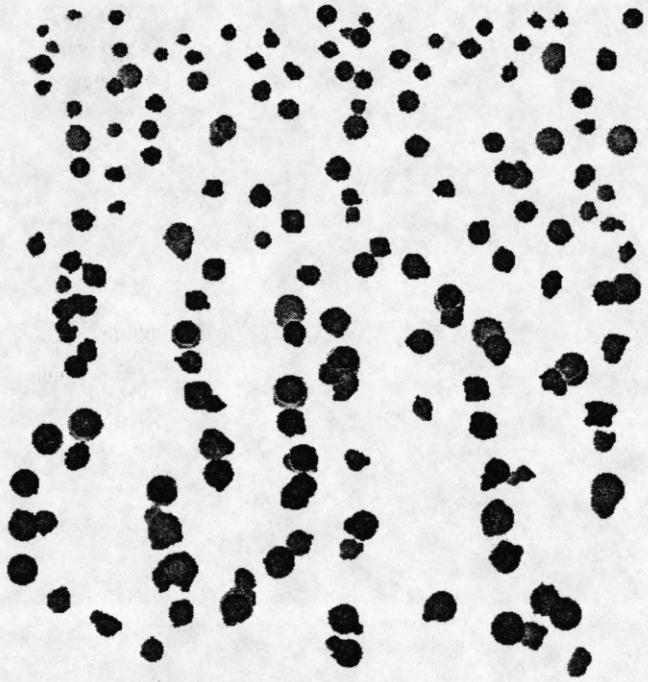
Figure 14, continued (Fleecy clouds; negative-contrast regions)
(e) Synthetic image to illustrate the planar fit A_c 160, slant 55°, tilt 280°. (f) and (g) Ratings of various possible planar fits. In (f) slant and tilt are varied while A_c is constant at 160. In (g) A_c is varied while slant and tilt are constant at 55° and 280° respectively.



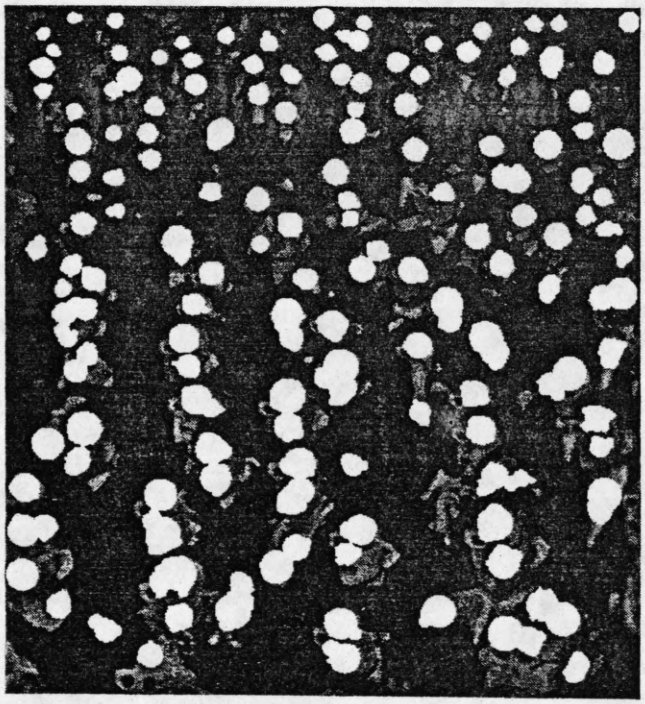
(a)



(b)



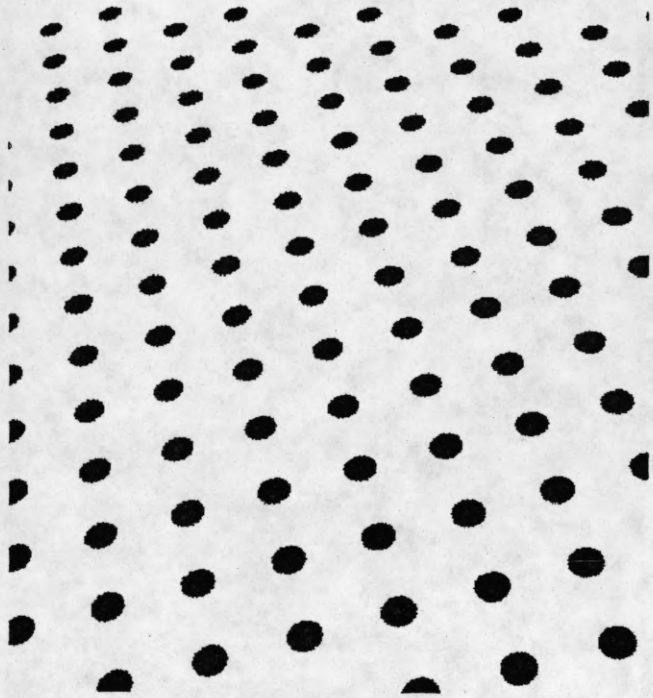
(c)



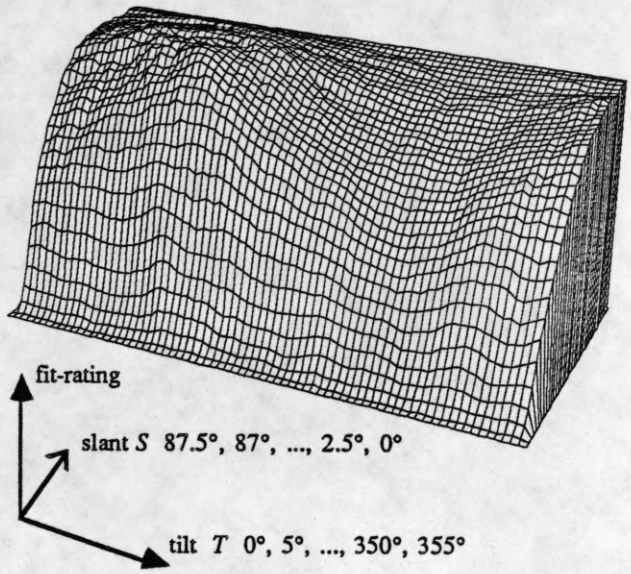
(d)

Figure 15

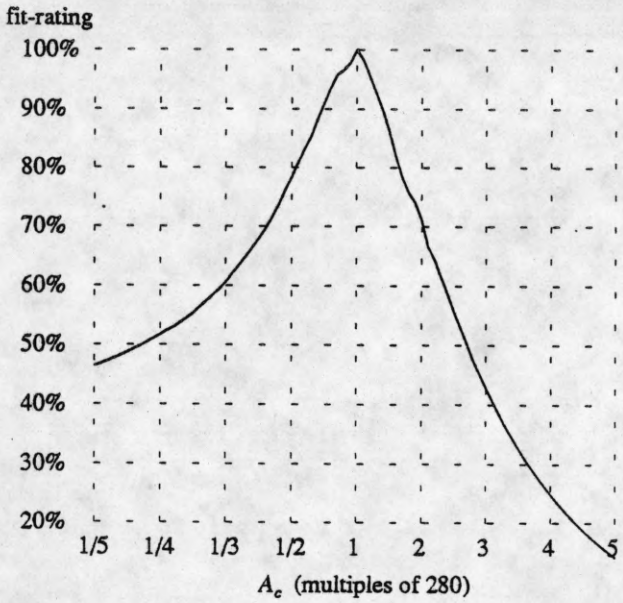
(a) Audience at a 3D movie. (b) Disks corresponding to positive-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. (c) Extracted texels. These are all regions (sets of overlapping disks) having area within a factor of two of the area expected by the best planar fit (A_c 280, slant 45°, tilt 105°). The texels that fit the plane most closely are printed darkest. (d) The texels superimposed on a dark reproduction of the original.



(e)

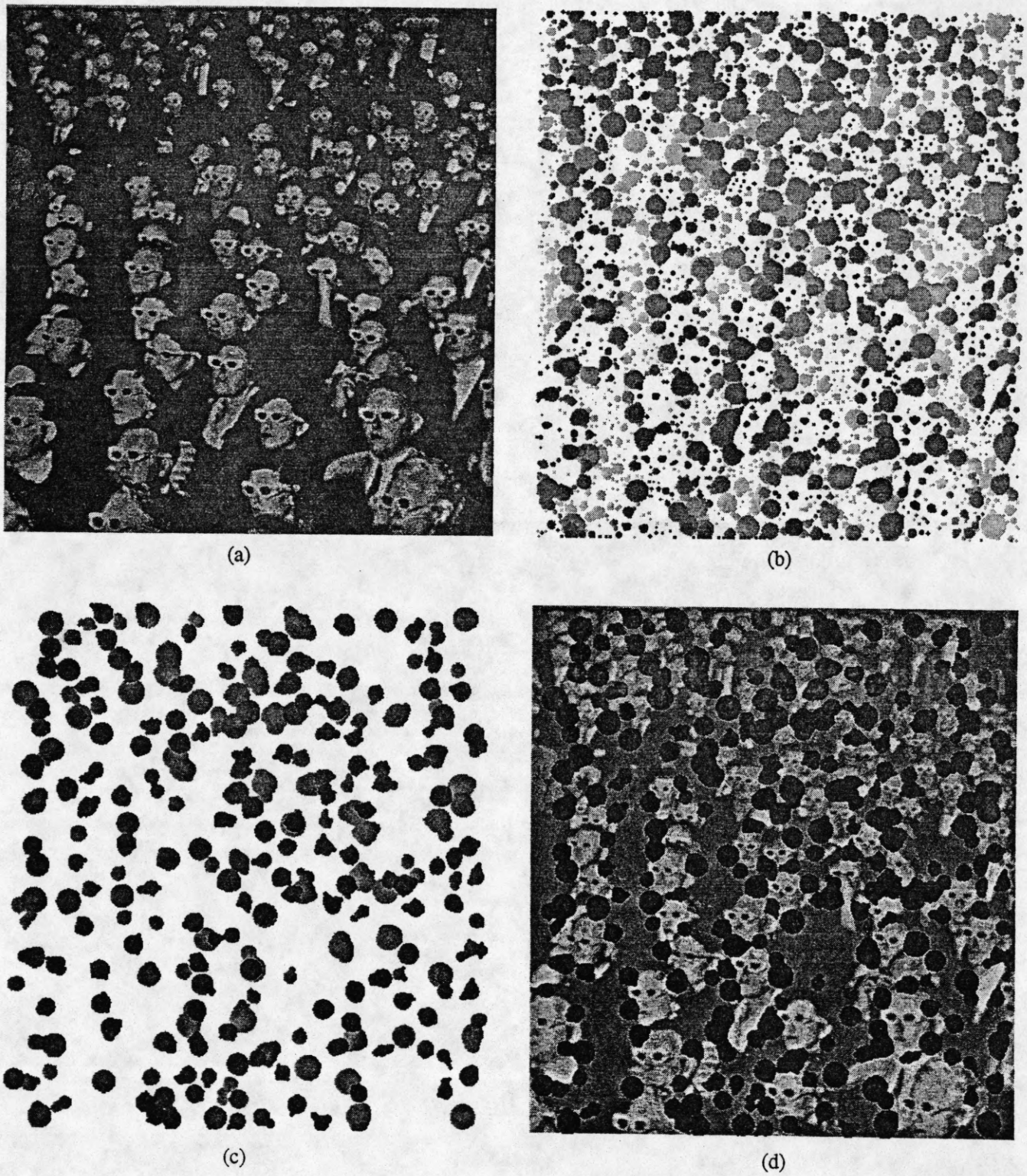


(f)



(g)

Figure 15, continued (Audience at a 3D movie; positive-contrast regions)
(e) Synthetic image to illustrate the planar fit A_c 280, slant 45°, tilt 105°. (f) and (g) Ratings of various possible planar fits. In (f) slant and tilt are varied while A_c is constant at 280. In (g) A_c is varied while slant and tilt are constant at 45° and 105° respectively.

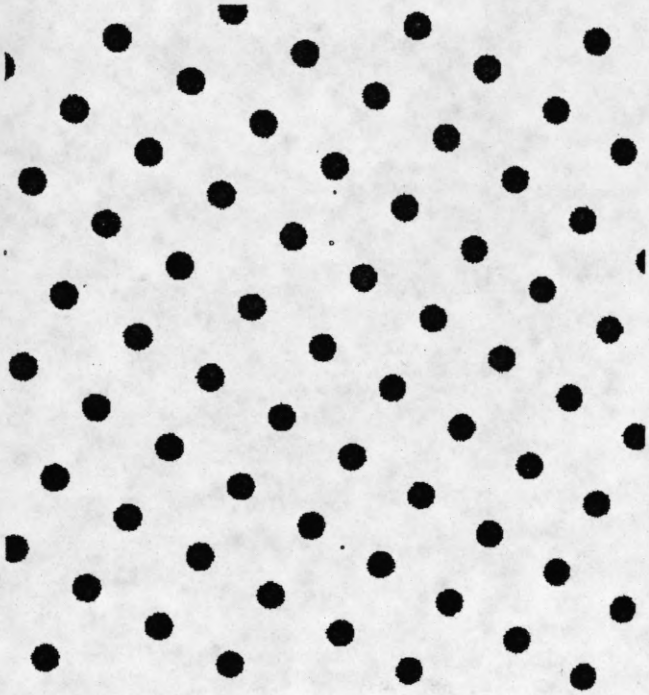


(a) (b)

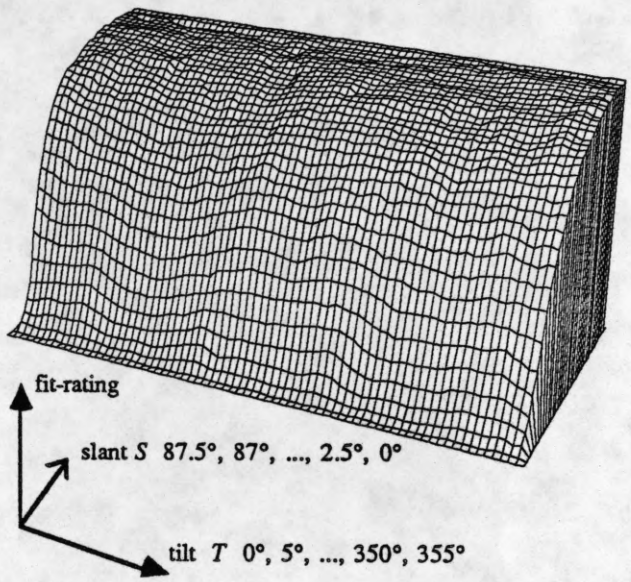
(c) (d)

Figure 16

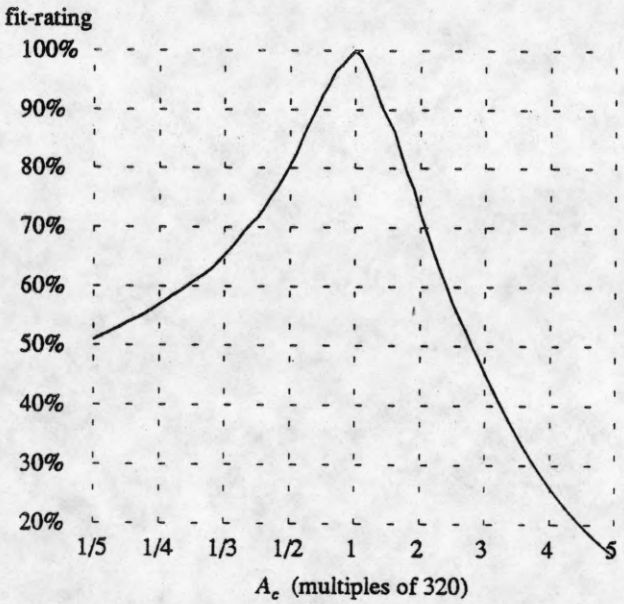
(a) Audience at a 3D movie. (b) Disks corresponding to negative-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. (c) Extracted texels. These are all regions (sets of overlapping disks) having area within a factor of two of the area expected by the best planar fit (A_e 320, slant 7.5°, tilt 330°). The texels that fit the plane most closely are printed darkest. (d) The texels superimposed on a bright reproduction of the original.



(e)

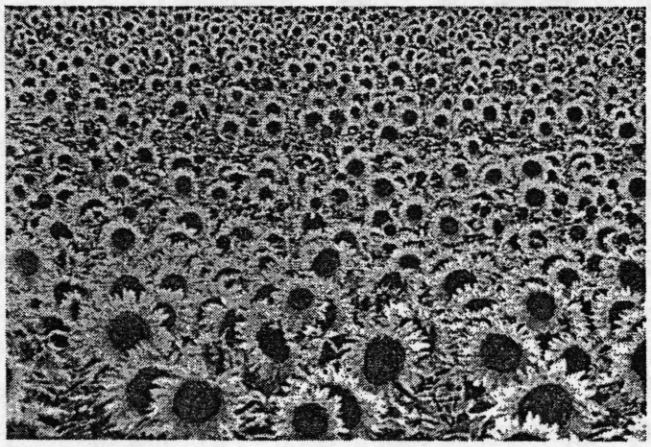


(f)

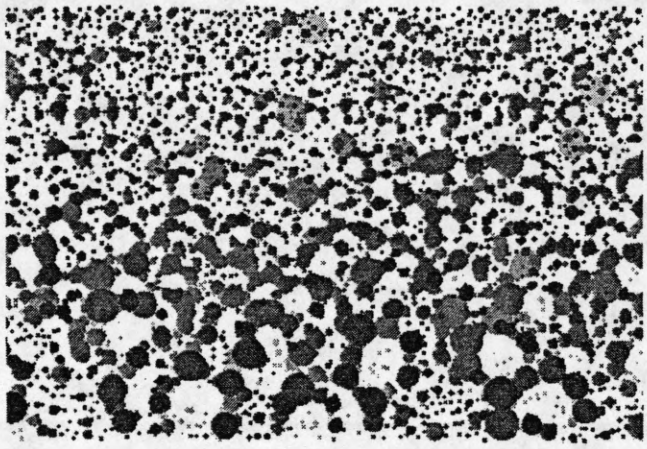


(g)

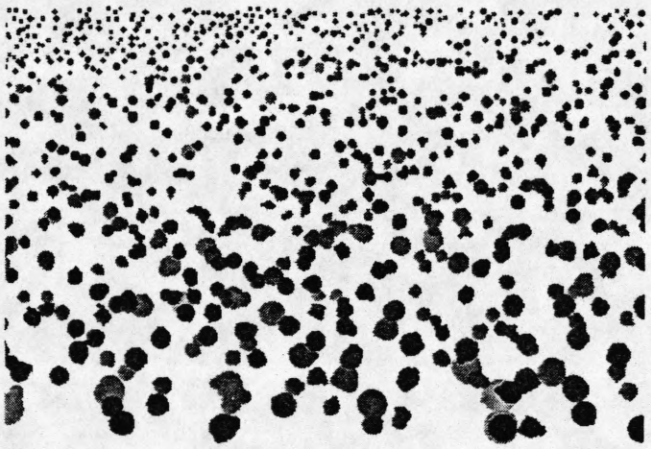
Figure 16, continued (Audience at a 3D movie; negative-contrast regions)
(e) Synthetic image to illustrate the planar fit A_c 320, slant 7.5°, tilt 330°. (f) and (g) Ratings of various possible planar fits. In (f) slant and tilt are varied while A_c is constant at 320. In (g) A_c is varied while slant and tilt are constant at 7.5° and 330° respectively.



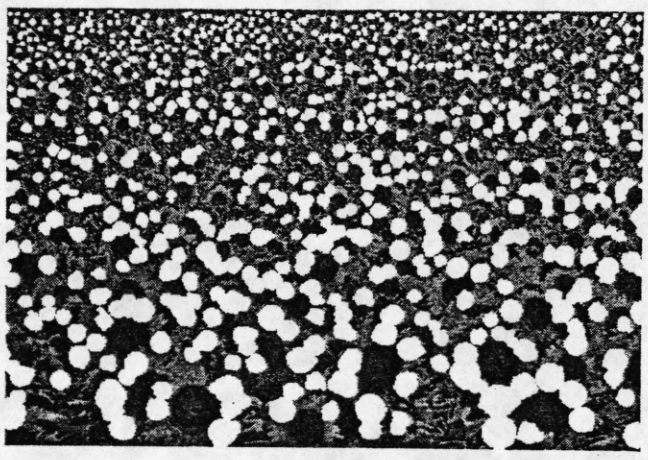
(a)



(b)



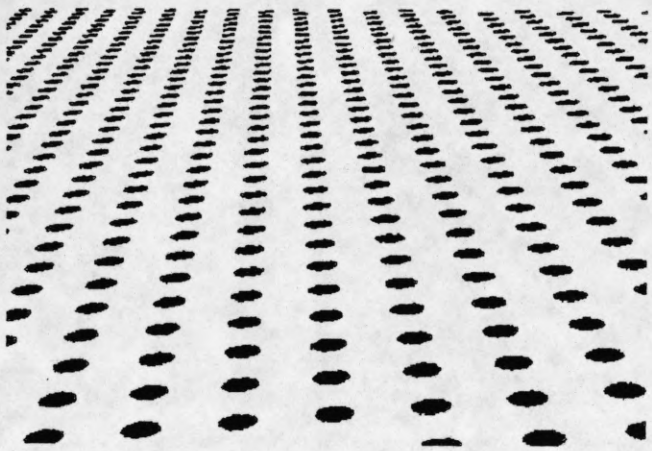
(c)



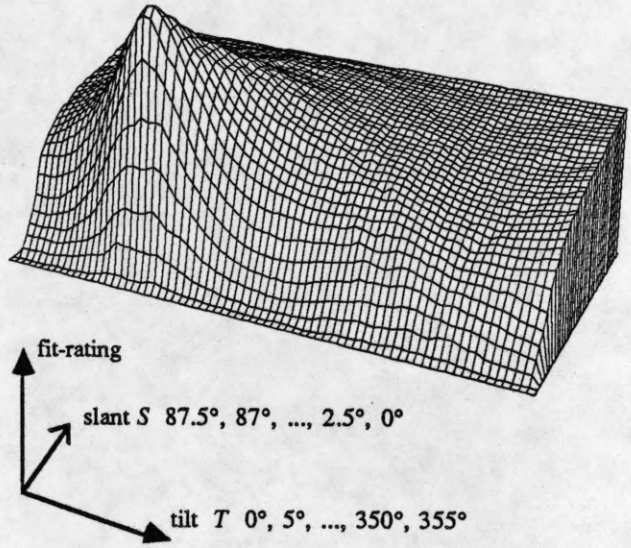
(d)

Figure 17

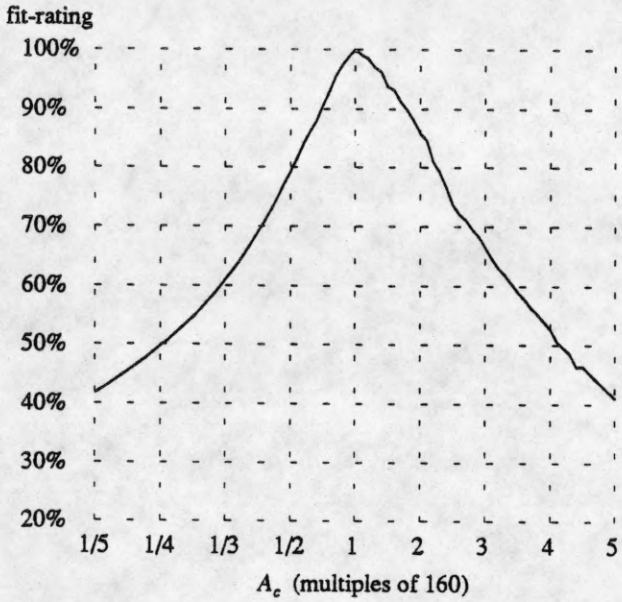
(a) Sunflowers. (b) Disks corresponding to positive-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. (c) Extracted texels. These are all regions (sets of overlapping disks) having area within a factor of two of the area expected by the best planar fit (A_c 160, slant 70°, tilt 95°). The texels that fit the plane most closely are printed darkest. (d) The texels superimposed on a dark reproduction of the original.



(e)

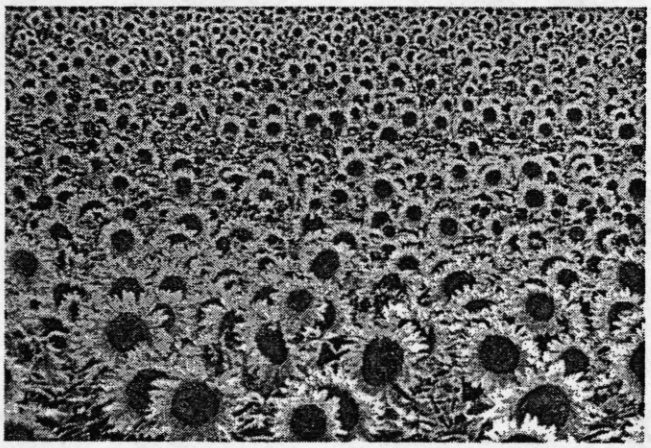


(f)

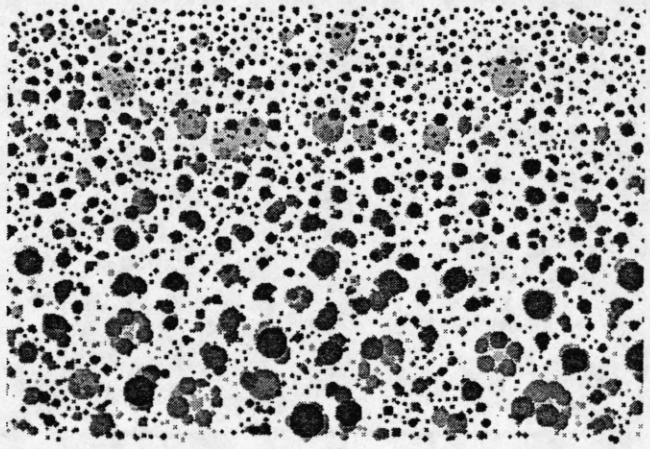


(g)

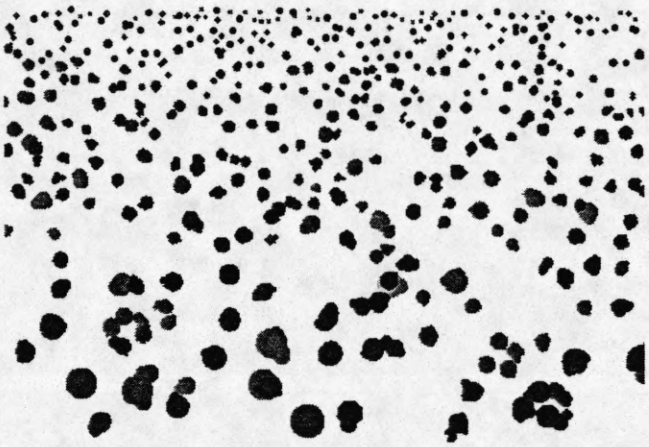
Figure 17, continued (Sunflowers; positive-contrast regions)
(e) Synthetic image to illustrate the planar fit A_c 160, slant 70°, tilt 95°. (f) and (g) Ratings of various possible planar fits. In (f) slant and tilt are varied while A_c is constant at 160. In (g) A_c is varied while slant and tilt are constant at 70° and 95° respectively.



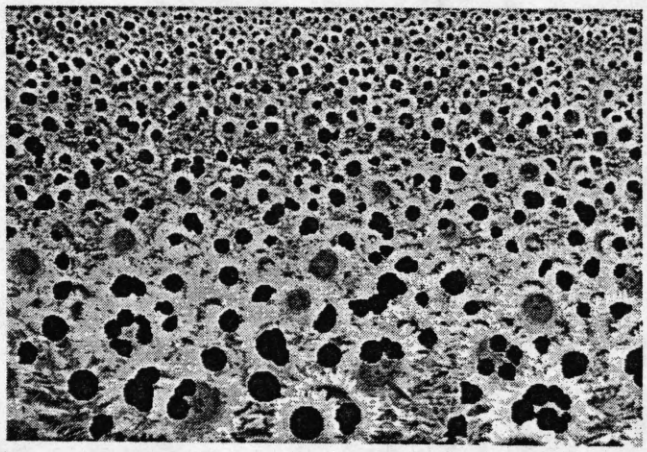
(a)



(b)



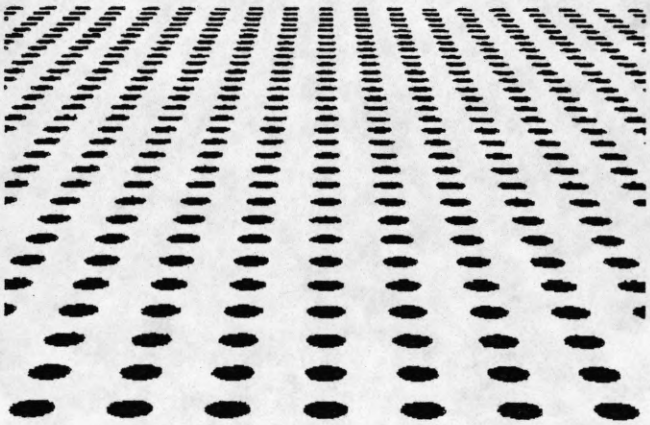
(c)



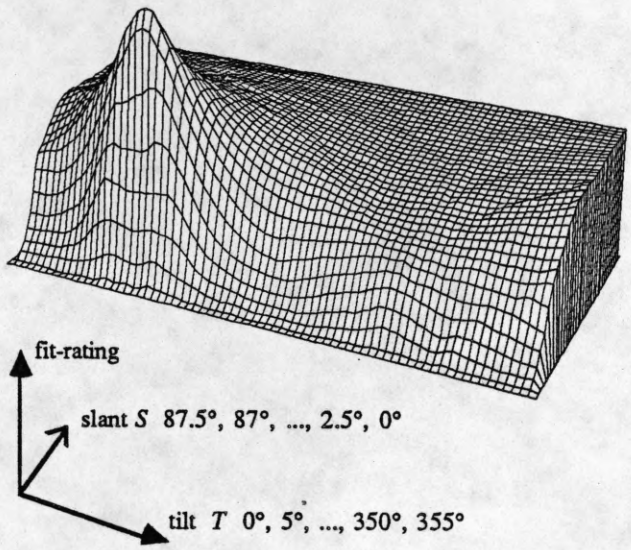
(d)

Figure 18

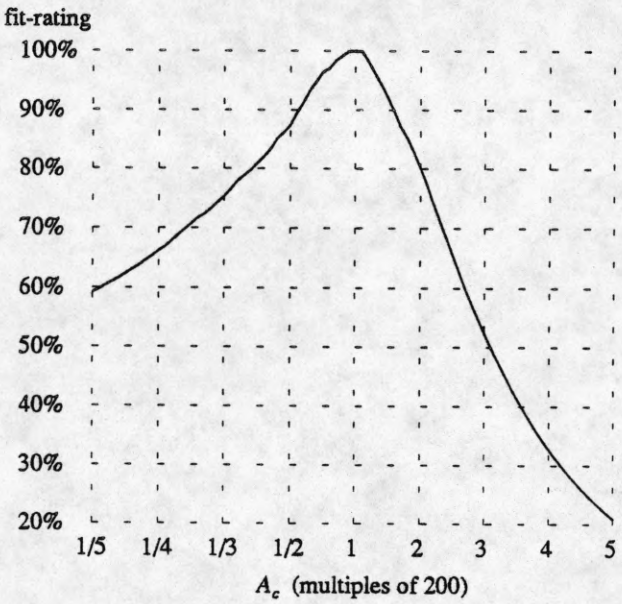
(a) Sunflowers. (b) Disks corresponding to negative-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. (c) Extracted texels. These are all regions (sets of overlapping disks) having area within a factor of two of the area expected by the best planar fit (A_e 200, slant 70°, tilt 90°). The texels that fit the plane most closely are printed darkest. (d) The texels superimposed on a bright reproduction of the original.



(e)

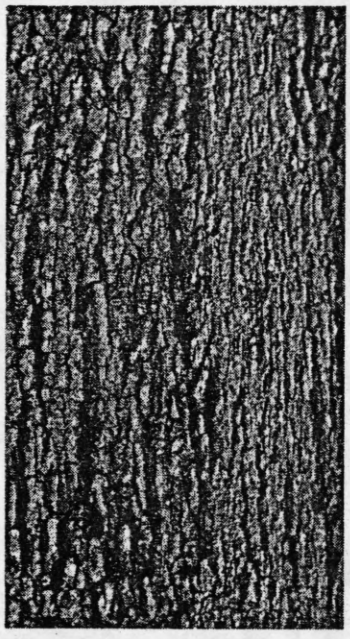


(f)

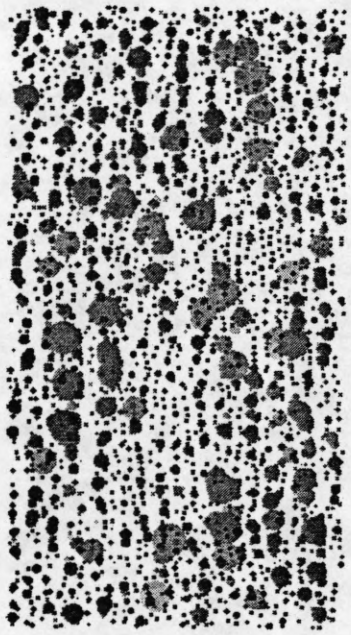


(g)

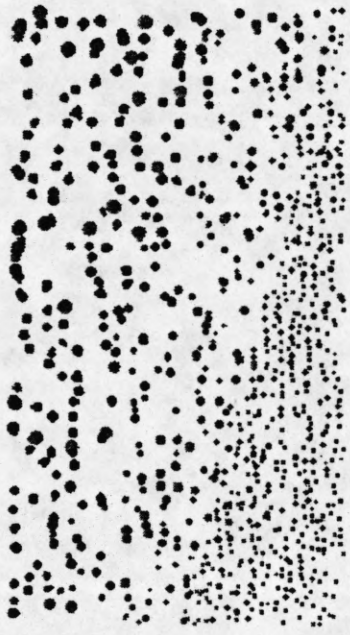
Figure 18, continued (Sunflowers; negative-contrast regions)
(e) Synthetic image to illustrate the planar fit A_c 200, slant 70°, tilt 90°. (f) and (g) Ratings of various possible planar fits. In (f) slant and tilt are varied while A_c is constant at 200. In (g) A_c is varied while slant and tilt are constant at 70° and 90° respectively.



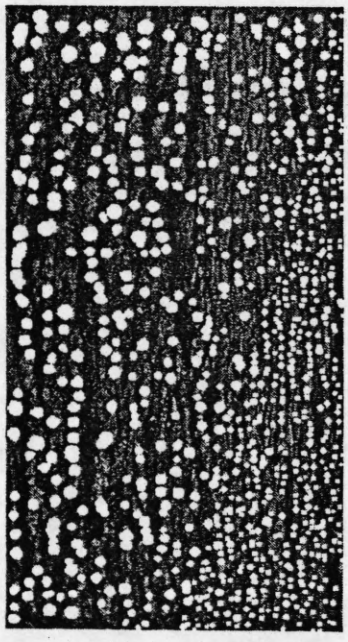
(a)



(b)



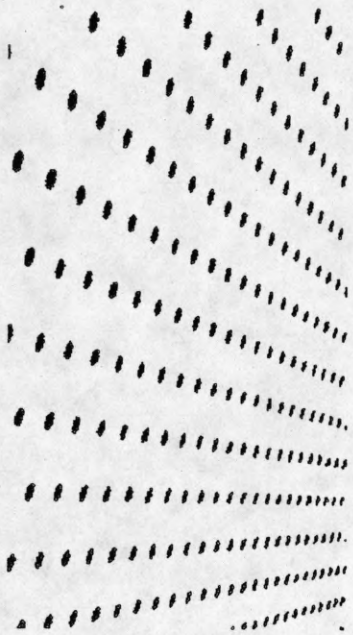
(c)



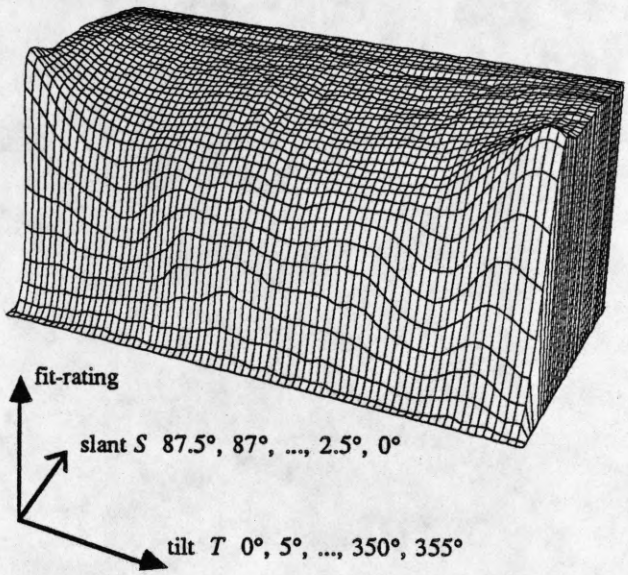
(d)

Figure 19

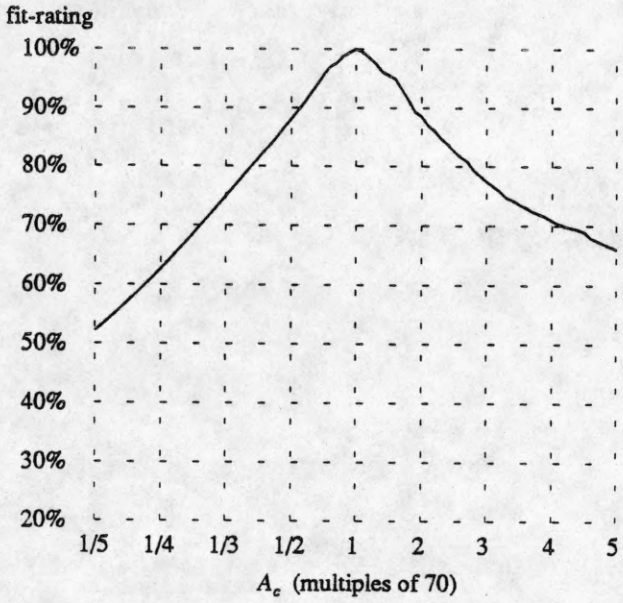
(a) Tree trunk. (b) Disks corresponding to positive-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. (c) Extracted texels. These are all regions (sets of overlapping disks) having area within a factor of two of the area expected by the best planar fit (A_c 70, slant 65°, tilt 345°). The texels that fit the plane most closely are printed darkest. (d) The texels superimposed on a dark reproduction of the original.



(e)

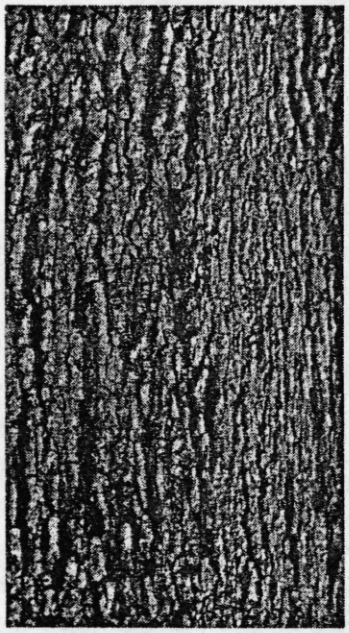


(f)

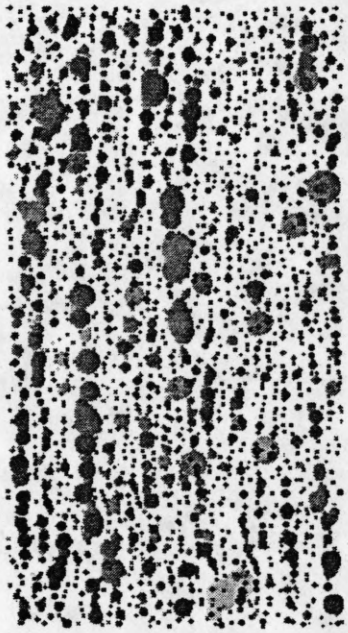


(g)

Figure 19, continued (Tree trunk; positive-contrast regions)
(e) Synthetic image to illustrate the planar fit A_c 70, slant 65°, tilt 345°. (f) and (g) Ratings of various possible planar fits. In (f) slant and tilt are varied while A_c is constant at 70. In (g) A_c is varied while slant and tilt are constant at 65° and 345° respectively.

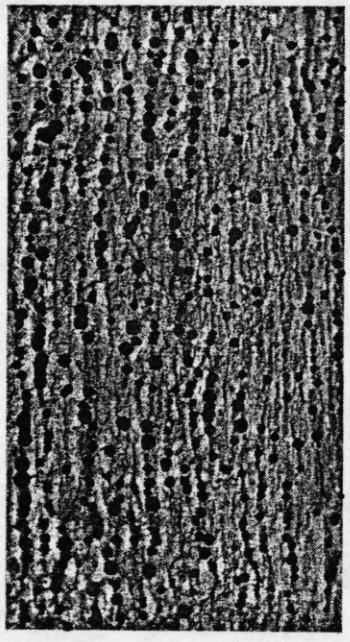


(a)



(b)

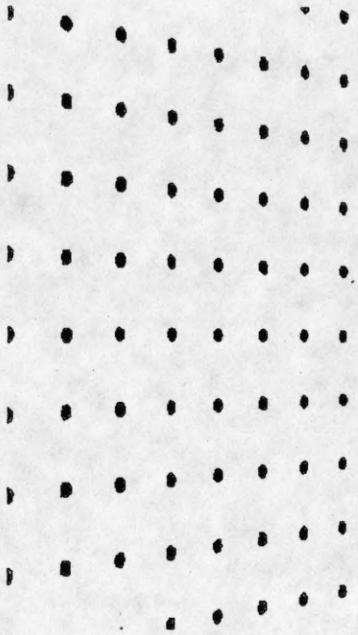
(c)



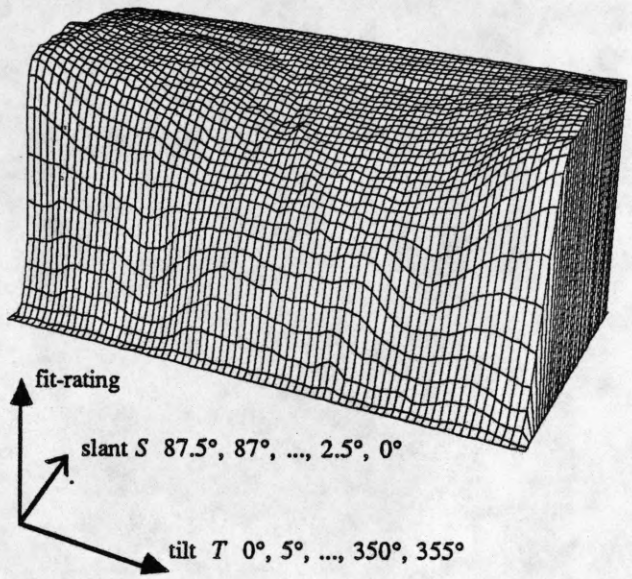
(d)

Figure 20

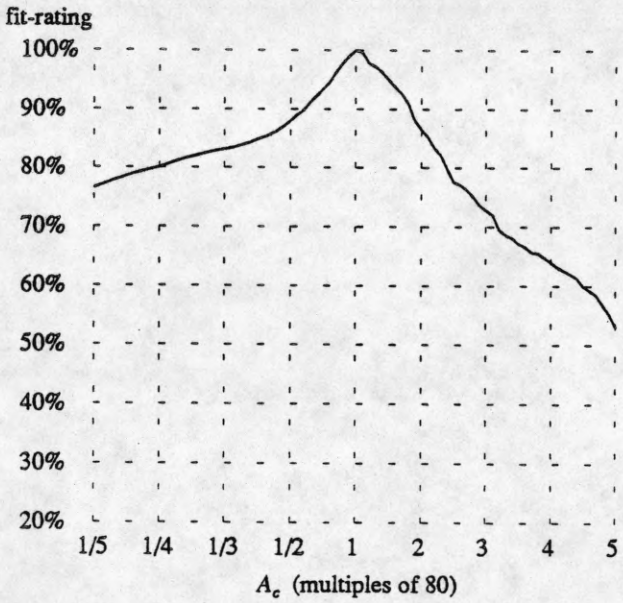
(a) Tree trunk. (b) Disks corresponding to negative-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. (c) Extracted texels. These are all regions (sets of overlapping disks) having area within a factor of two of the area expected by the best planar fit (A_c 80, slant 42.5°, tilt 0°). The texels that fit the plane most closely are printed darkest. (d) The texels superimposed on a bright reproduction of the original.



(e)

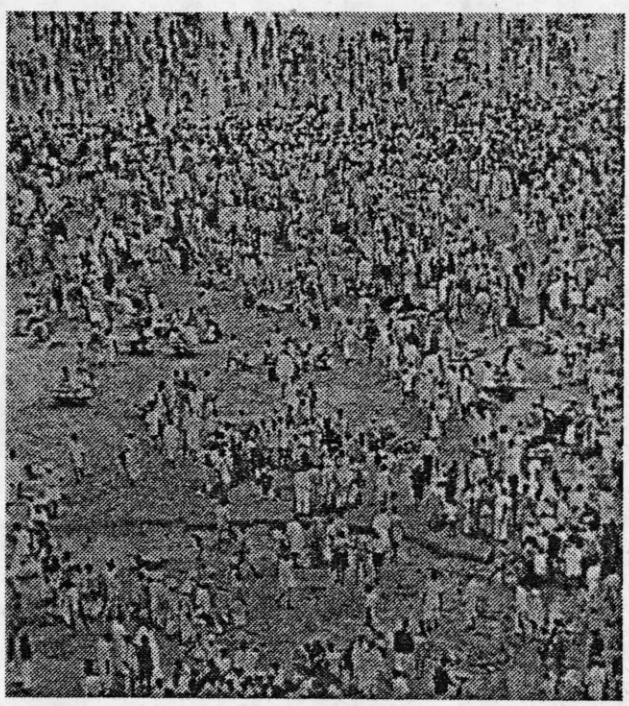


(f)

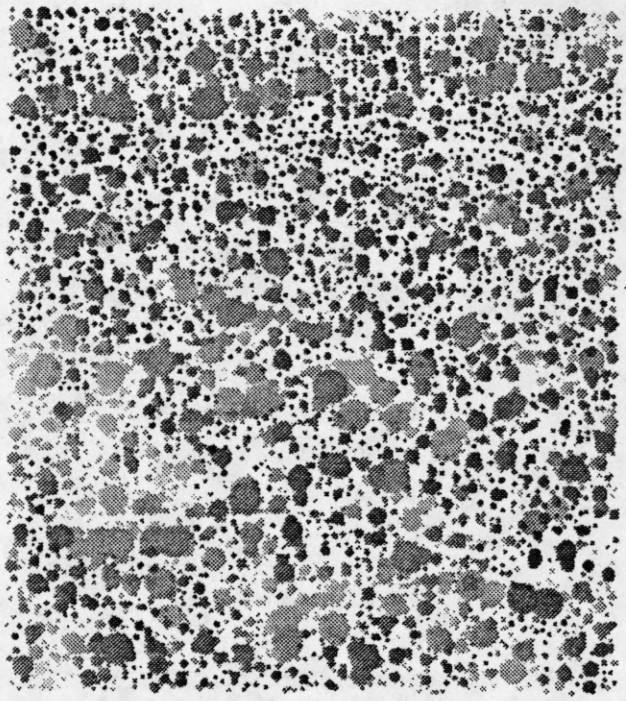


(g)

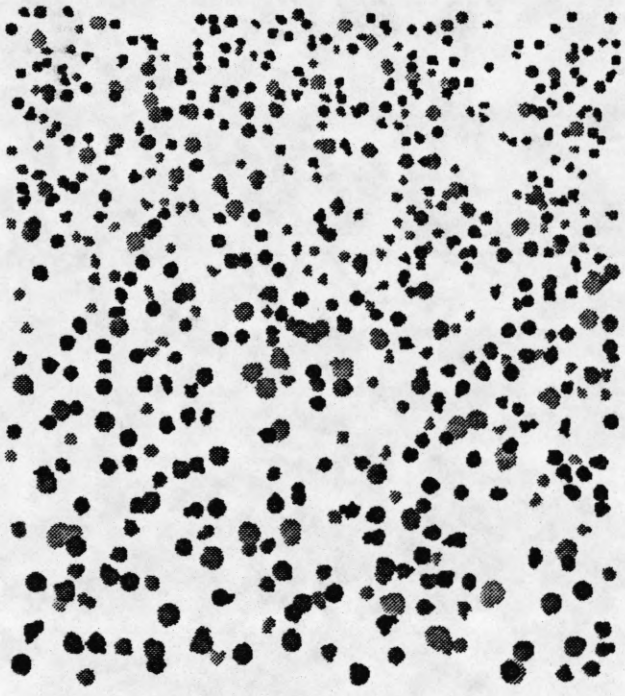
Figure 20, continued (Tree trunk; negative-contrast regions)
(e) Synthetic image to illustrate the planar fit A_c 80, slant 42.5°, tilt 0°. (f) and (g) Ratings of various possible planar fits. In (f) slant and tilt are varied while A_c is constant at 80. In (g) A_c is varied while slant and tilt are constant at 42.5° and 0° respectively.



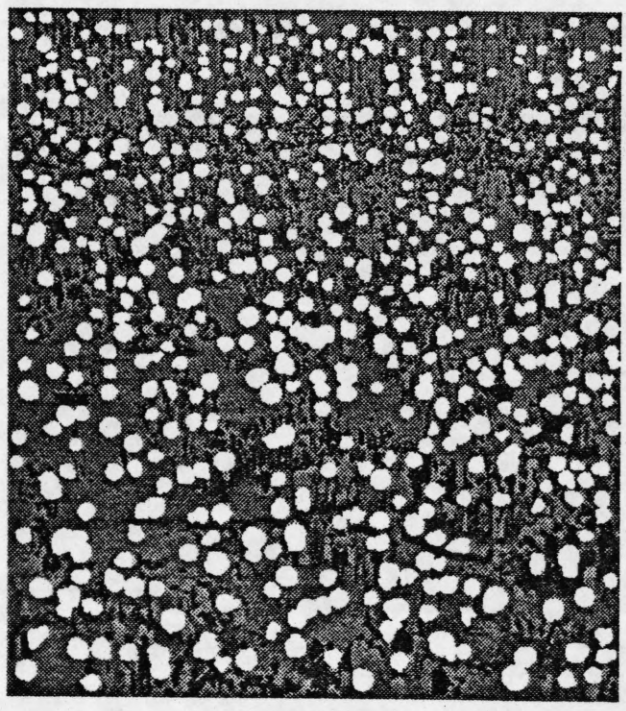
(a)



(b)



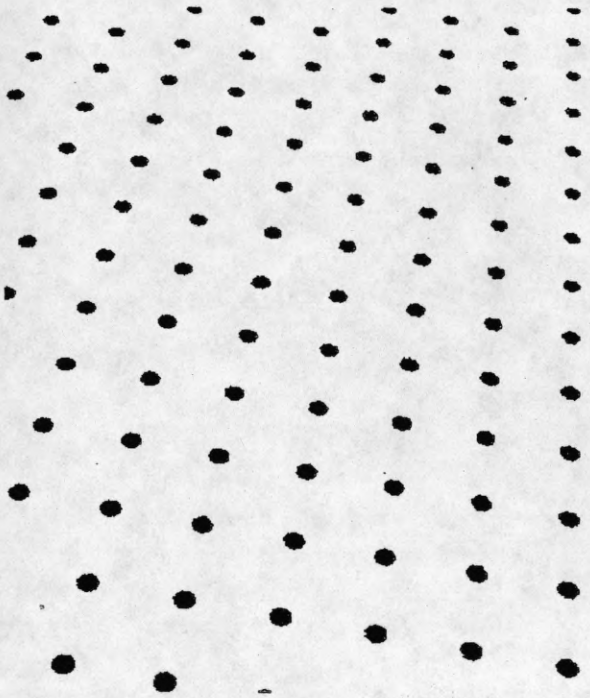
(c)



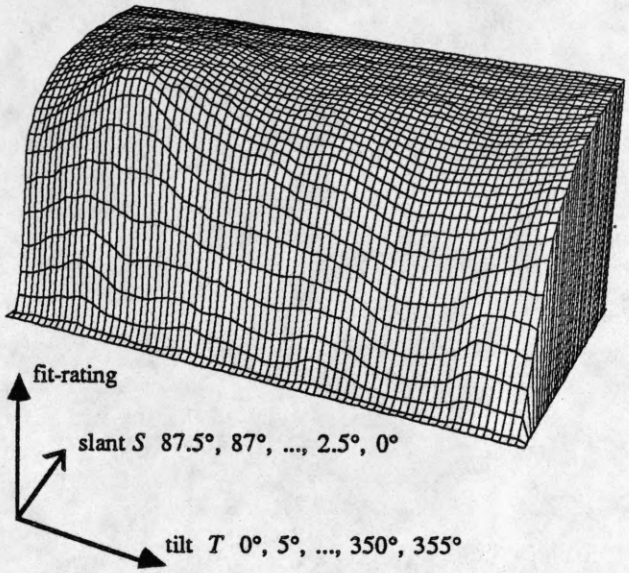
(d)

Figure 21

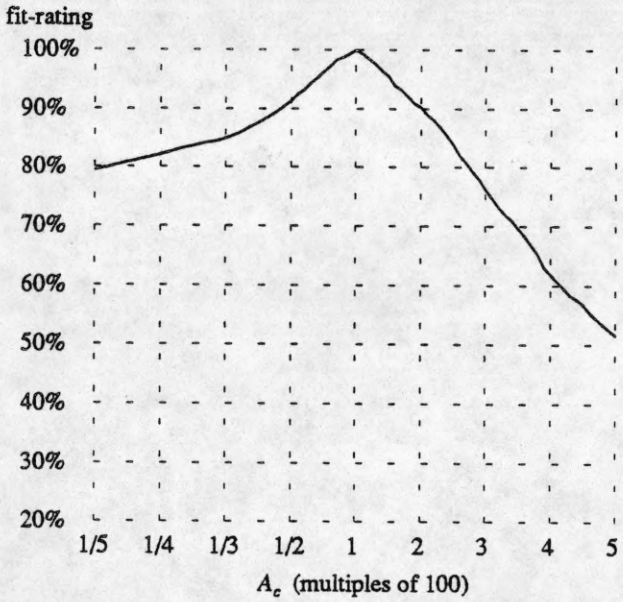
(a) Bathers on the Ganges. (b) Disks corresponding to positive-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. (c) Extracted texels. These are all regions (sets of overlapping disks) having area within a factor of two of the area expected by the best planar fit (A_c 100, slant 45°, tilt 80°). The texels that fit the plane most closely are printed darkest. (d) The texels superimposed on a dark reproduction of the original.



(e)



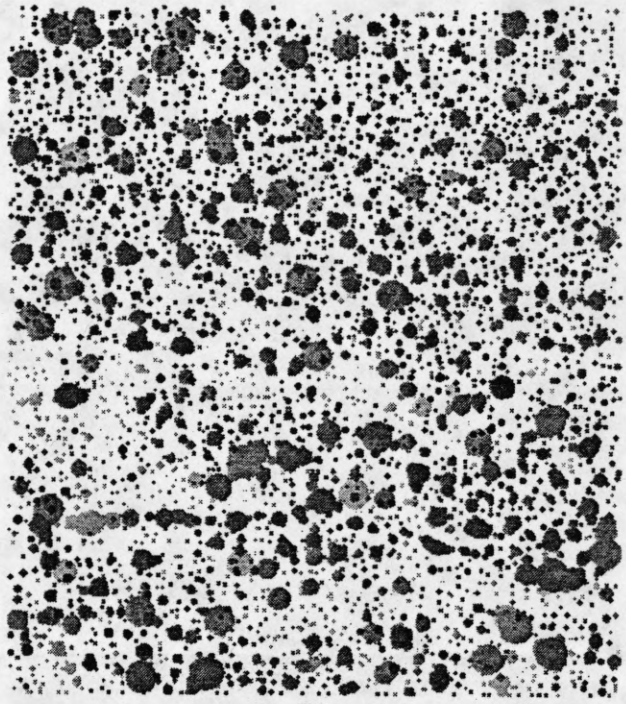
(f)



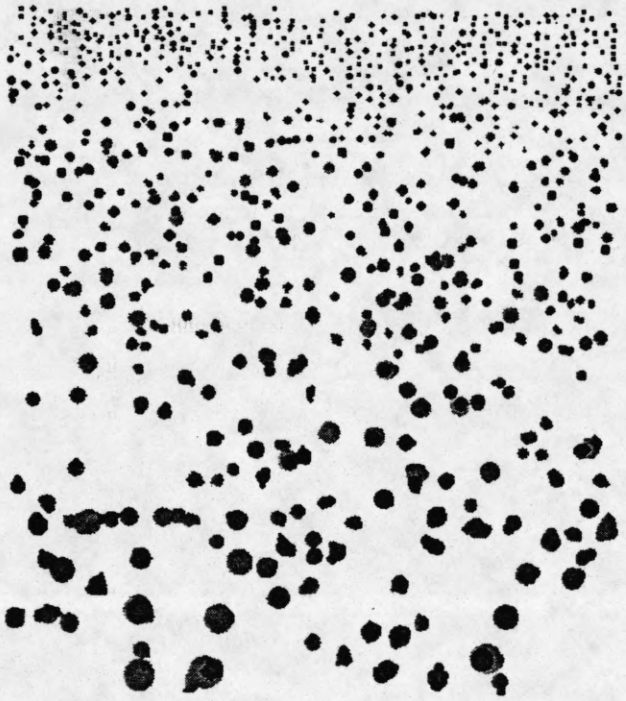
(g)

Figure 21, continued (Bathers on the Ganges; positive-contrast regions)
(e) Synthetic image to illustrate the planar fit A_c 100, slant 45°, tilt 80°. (f) and (g) Ratings of various possible planar fits. In (f) slant and tilt are varied while A_c is constant at 100. In (g) A_c is varied while slant and tilt are constant at 45° and 80° respectively.

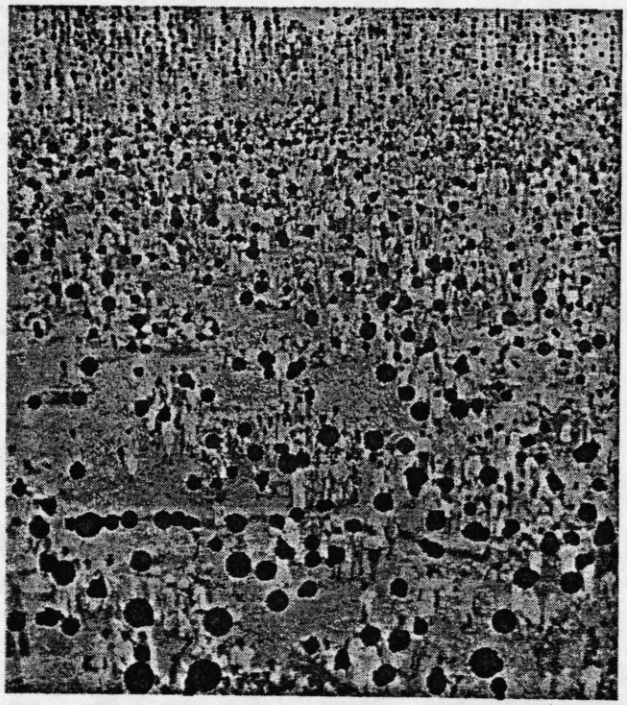
(a)



(b)



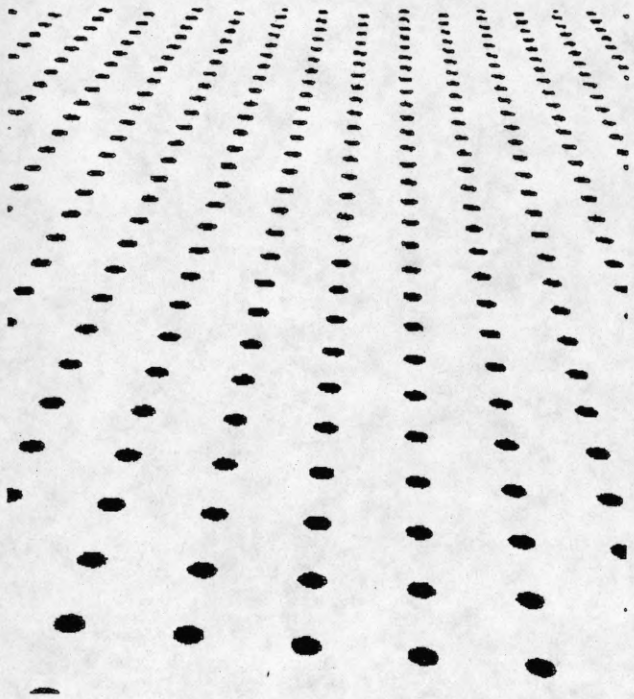
(c)



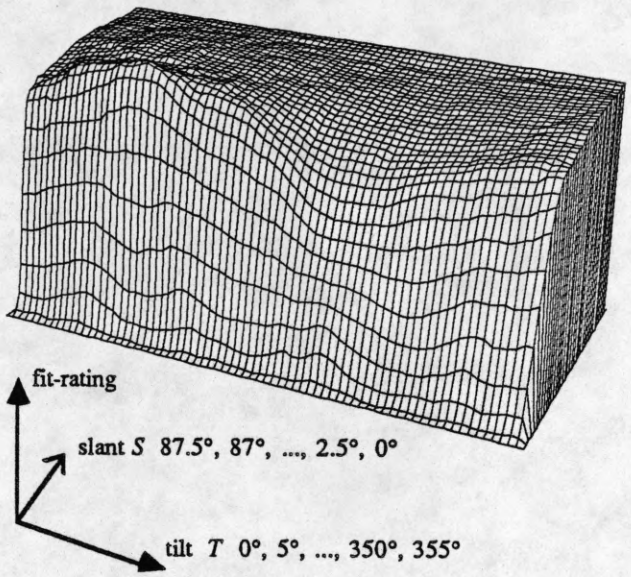
(d)

Figure 22

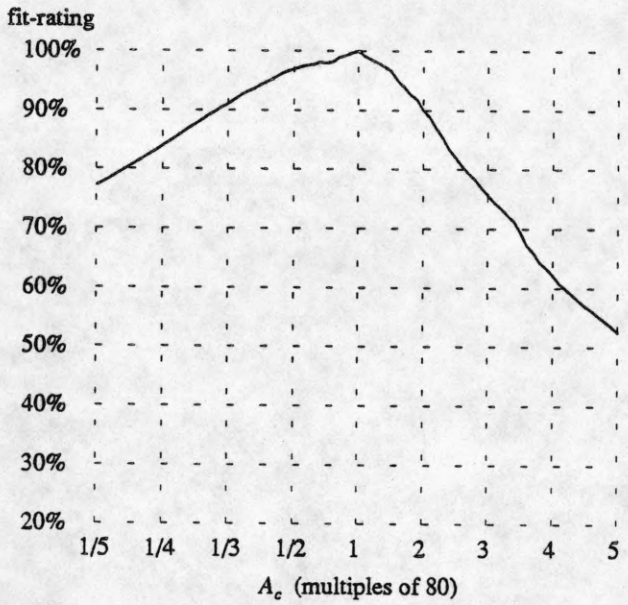
(a) Bathers on the Ganges. (b) Disks corresponding to negative-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. (c) Extracted texels. These are all regions (sets of overlapping disks) having area within a factor of two of the area expected by the best planar fit (A_c 80, slant 65°, tilt 85°). The texels that fit the plane most closely are printed darkest. (d) The texels superimposed on a bright reproduction of the original.



(e)



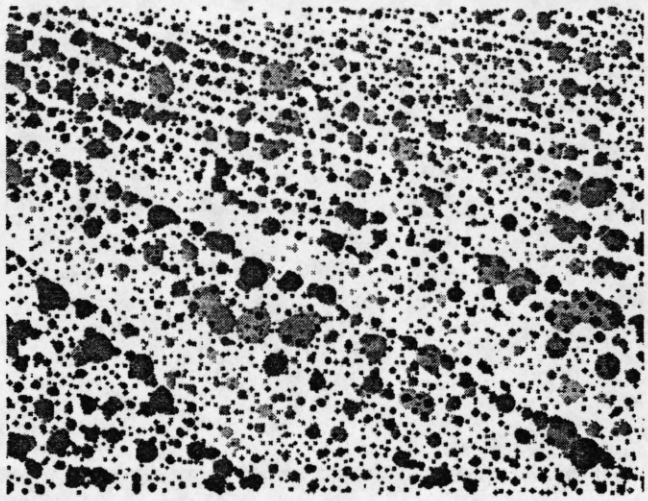
(f)



(g)

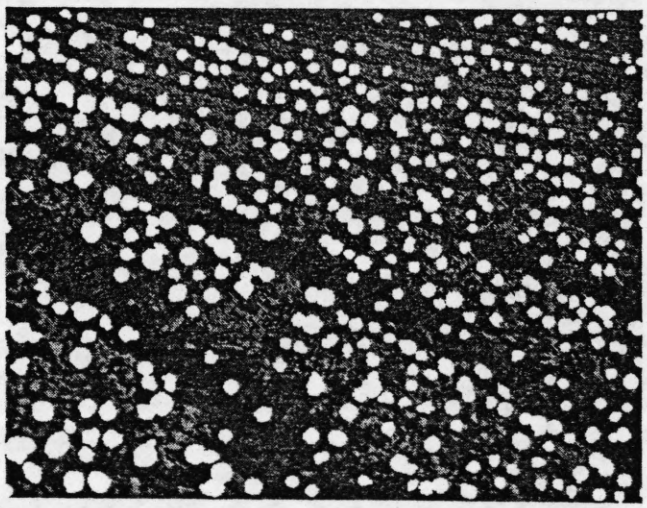
Figure 22, continued (Bathers on the Ganges; negative-contrast regions)
(e) Synthetic image to illustrate the planar fit A_c 80, slant 65°, tilt 85°. (f) and (g) Ratings of various possible planar fits. In (f) slant and tilt are varied while A_c is constant at 80. In (g) A_c is varied while slant and tilt are constant at 65° and 85° respectively.

(a)



(b)

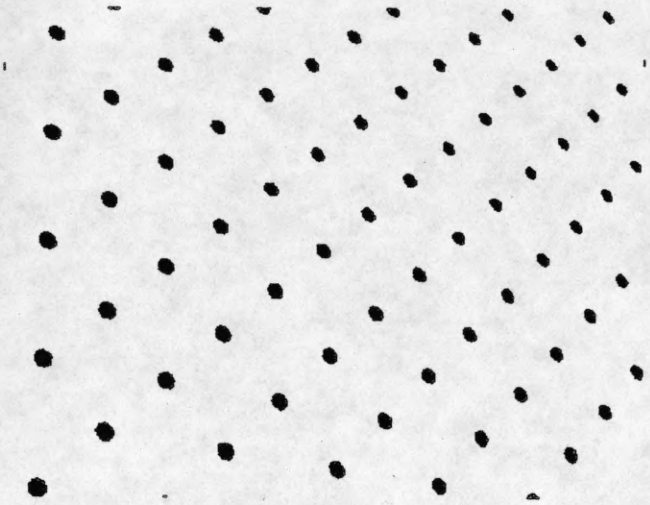
(c)



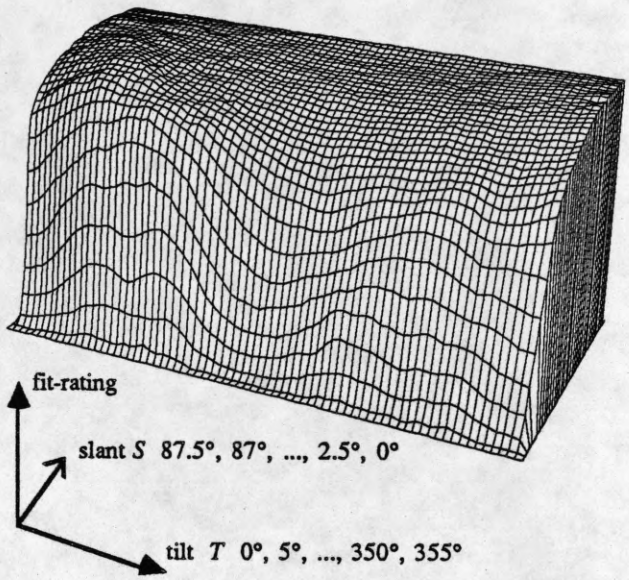
(d)

Figure 23

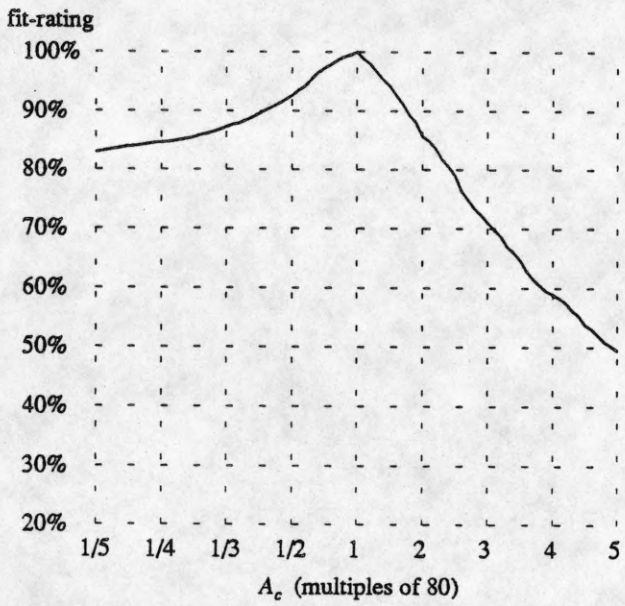
(a) A plowed field. (b) Disks corresponding to positive-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. (c) Extracted texels. These are all regions (sets of overlapping disks) having area within a factor of two of the area expected by the best planar fit (A_c 80, slant 42.5°, tilt 40°). The texels that fit the plane most closely are printed darkest. (d) The texels superimposed on a dark reproduction of the original.



(e)



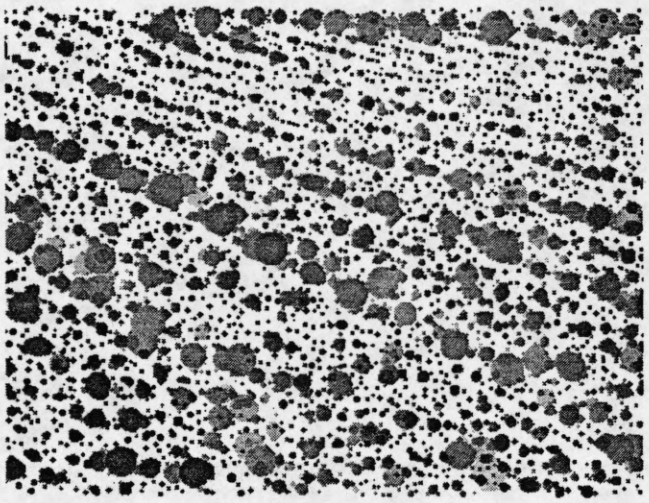
(f)



(g)

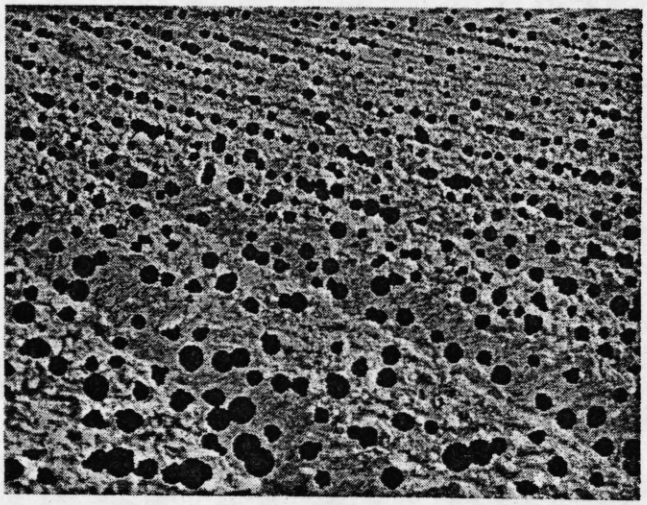
Figure 23, continued (A plowed field; positive-contrast regions)
(e) Synthetic image to illustrate the planar fit A_c 80, slant 42.5°, tilt 40°. (f) and (g) Ratings of various possible planar fits. In (f) slant and tilt are varied while A_c is constant at 80. In (g) A_c is varied while slant and tilt are constant at 42.5° and 40° respectively.

(a)



(b)

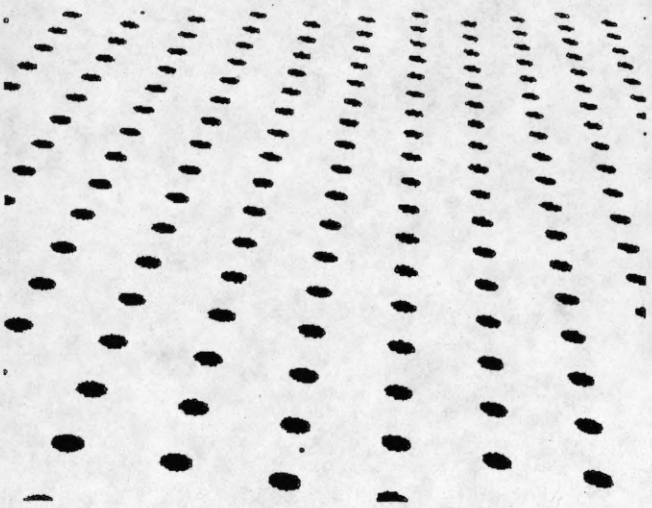
(c)



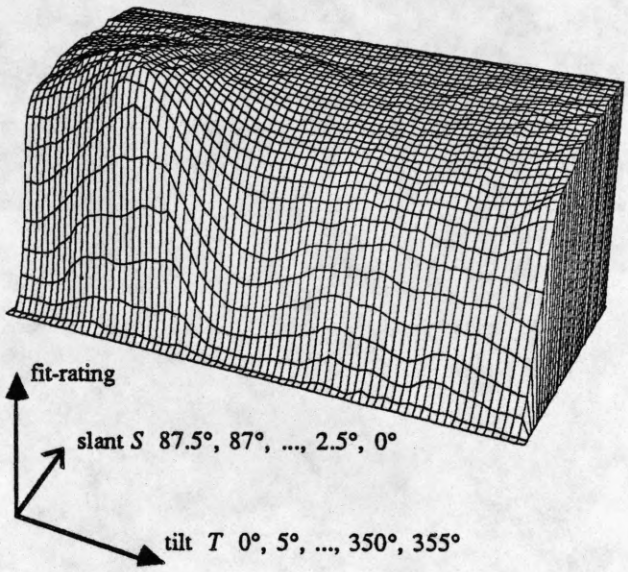
(d)

Figure 24

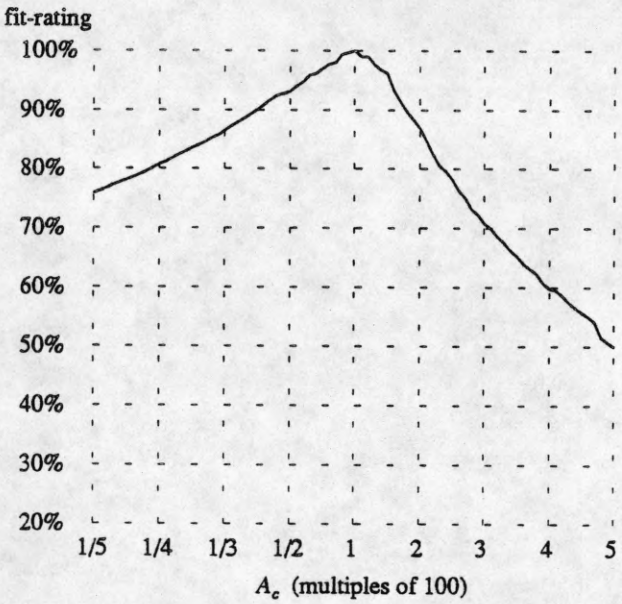
(a) A plowed field. (b) Disks corresponding to negative-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. (c) Extracted texels. These are all regions (sets of overlapping disks) having area within a factor of two of the area expected by the best planar fit (A_c 100, slant 65°, tilt 80°). The texels that fit the plane most closely are printed darkest. (d) The texels superimposed on a bright reproduction of the original.



(e)

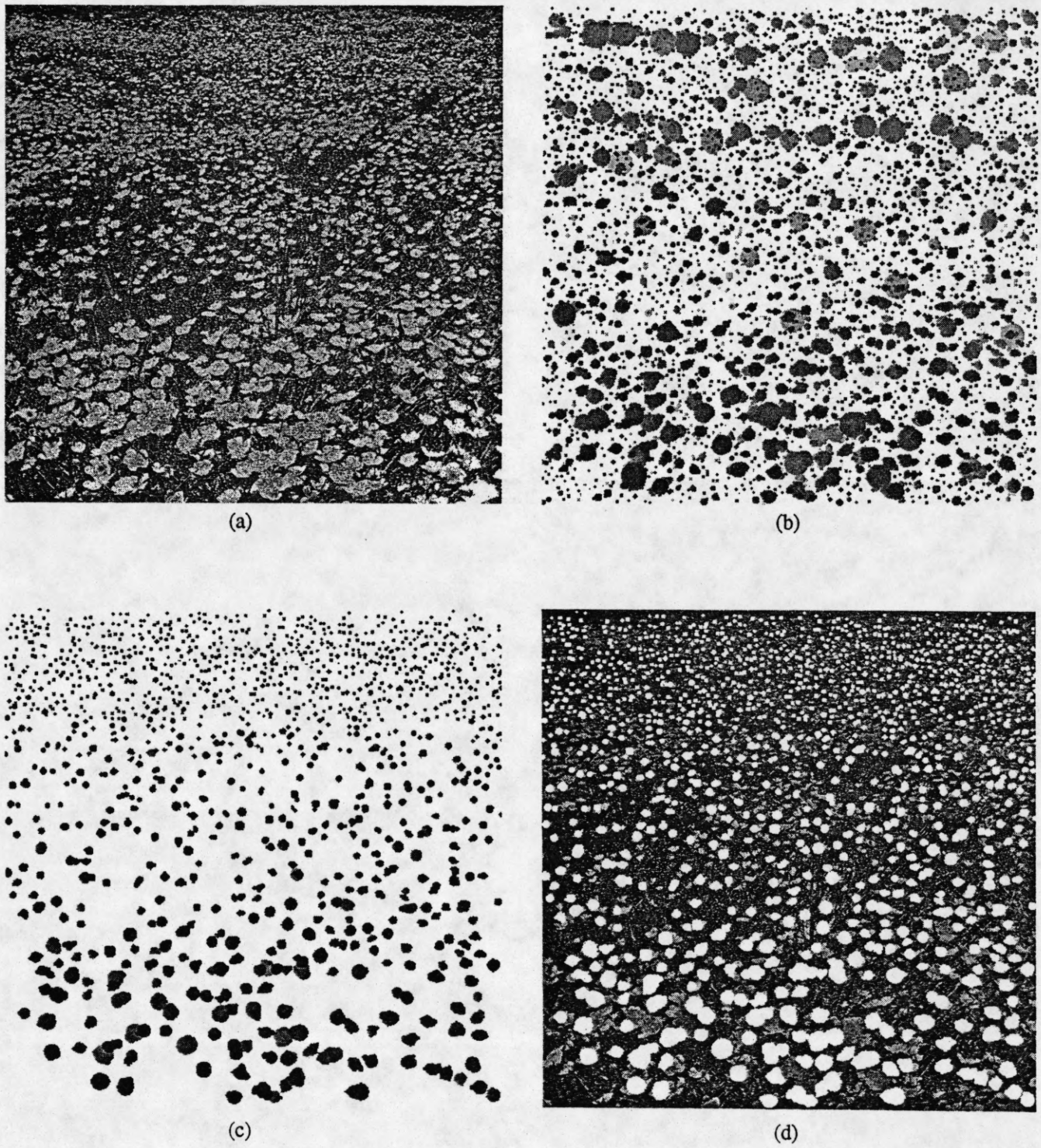


(f)



(g)

Figure 24, continued (A plowed field; negative-contrast regions)
(e) Synthetic image to illustrate the planar fit A_c 100, slant 65°, tilt 80°. (f) and (g) Ratings of various possible planar fits. In (f) slant and tilt are varied while A_c is constant at 100. In (g) A_c is varied while slant and tilt are constant at 65° and 80° respectively.

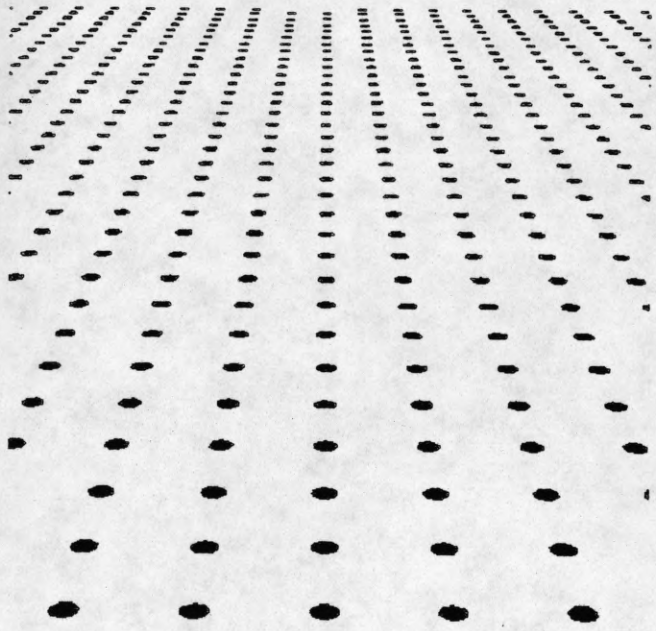


(a) (b)

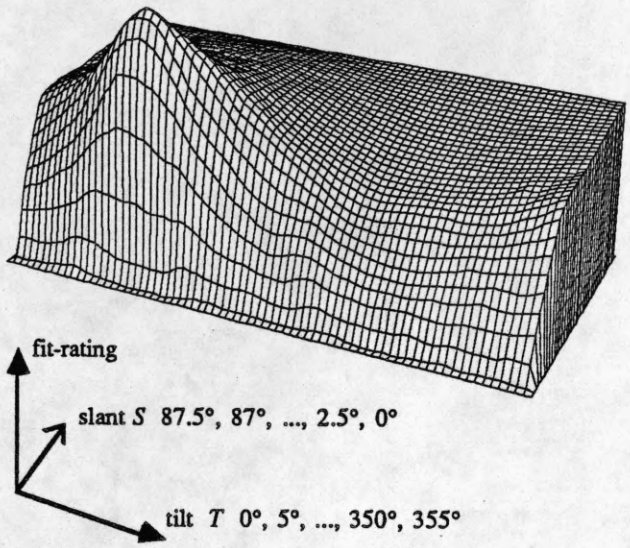
(c) (d)

Figure 25

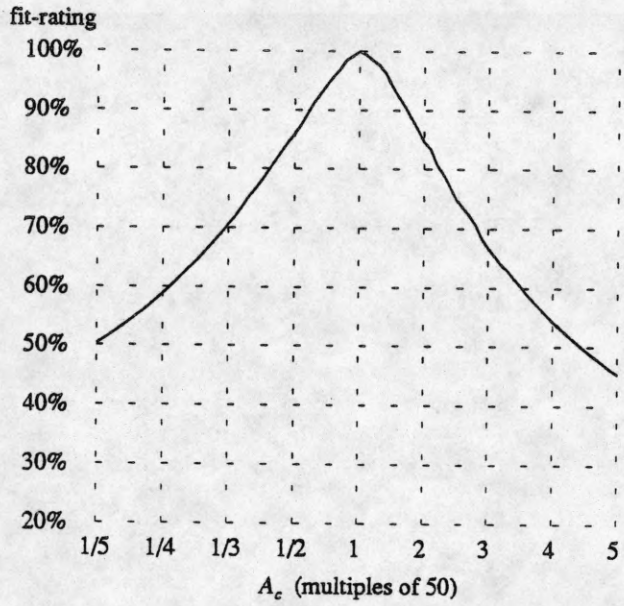
(a) A field of flowers. (b) Disks corresponding to positive-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. (c) Extracted texels. These are all regions (sets of overlapping disks) having area within a factor of two of the area expected by the best planar fit (A_c 50, slant 70°, tilt 90°). The texels that fit the plane most closely are printed darkest. (d) The texels superimposed on a dark reproduction of the original.



(e)



(f)



(g)

Figure 25, continued (A field of flowers; positive-contrast regions)
(e) Synthetic image to illustrate the planar fit A_c 50, slant 70°, tilt 90°. (f) and (g) Ratings of various possible planar fits. In (f) slant and tilt are varied while A_c is constant at 50. In (g) A_c is varied while slant and tilt are constant at 70° and 90° respectively.

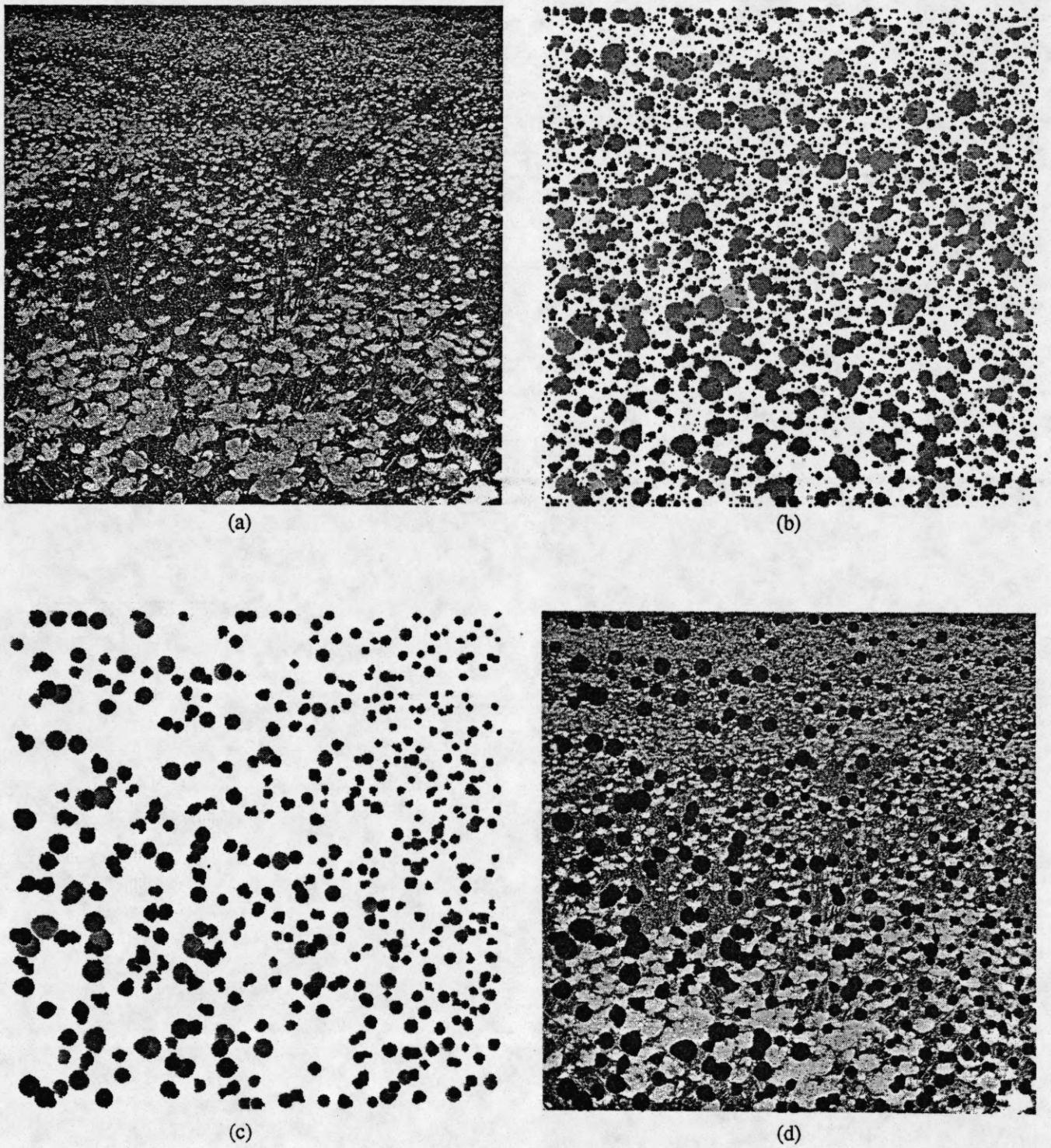
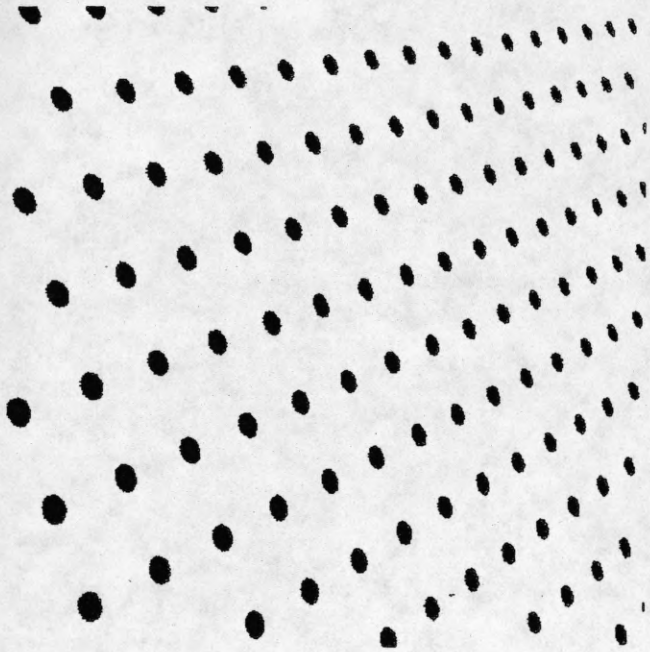
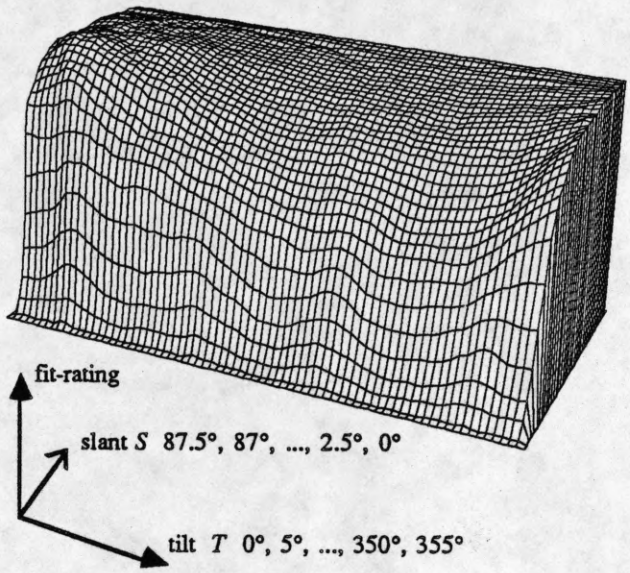


Figure 26

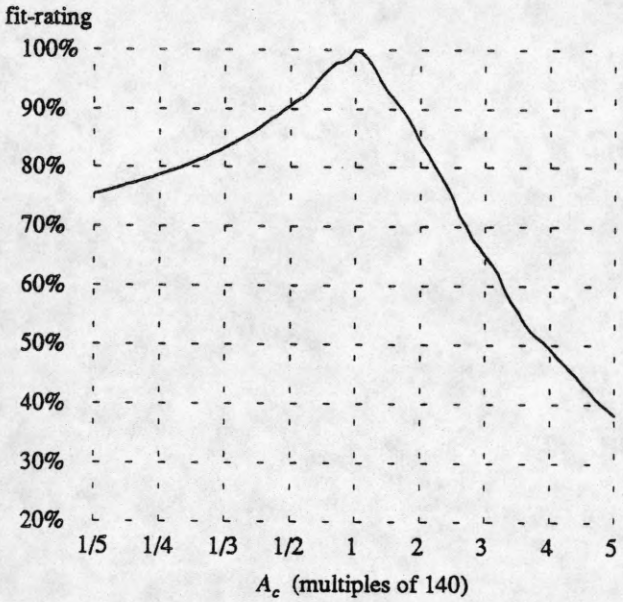
(a) A field of flowers. (b) Disks corresponding to negative-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. (c) Extracted texels. These are all regions (sets of overlapping disks) having area within a factor of two of the area expected by the best planar fit (A_c 140, slant 52.5°, tilt 20°). The texels that fit the plane most closely are printed darkest. (d) The texels superimposed on a bright reproduction of the original.



(e)



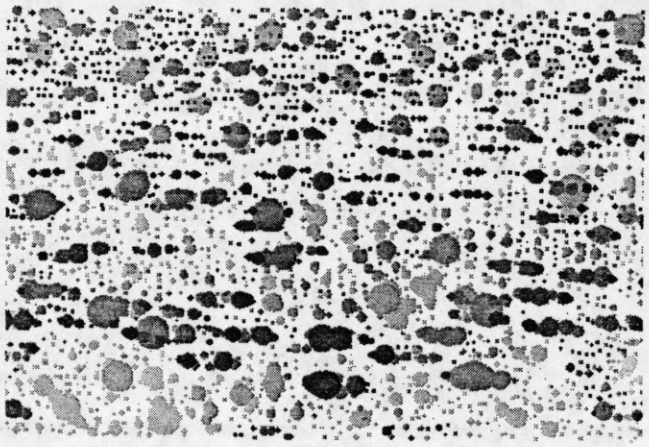
(f)



(g)

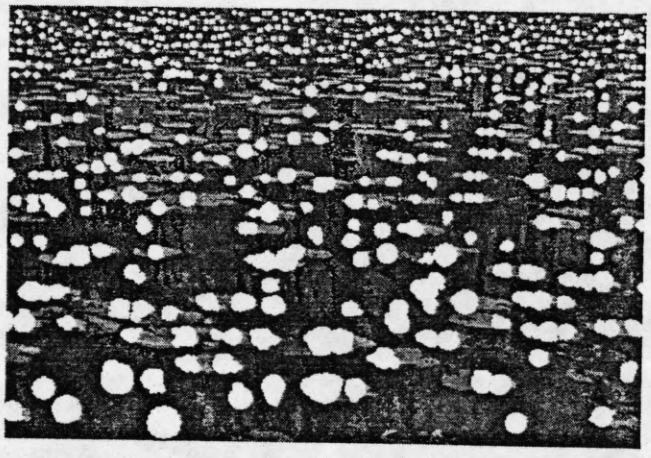
Figure 26, continued (A field of flowers; negative-contrast regions)
(e) Synthetic image to illustrate the planar fit A_c 140, slant 52.5°, tilt 20°. (f) and (g) Ratings of various possible planar fits. In (f) slant and tilt are varied while A_c is constant at 140. In (g) A_c is varied while slant and tilt are constant at 52.5° and 20° respectively.

(a)



(b)

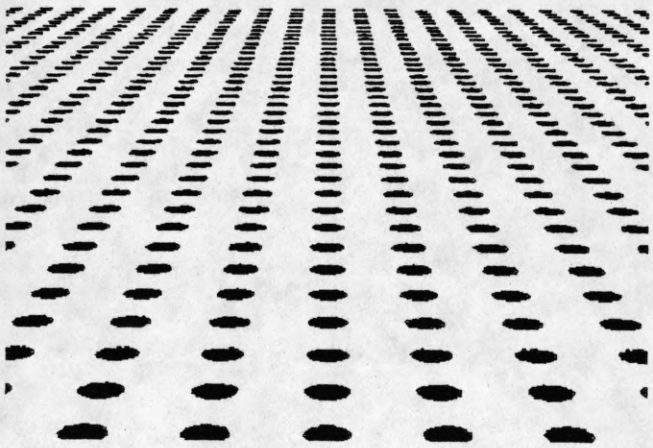
(c)



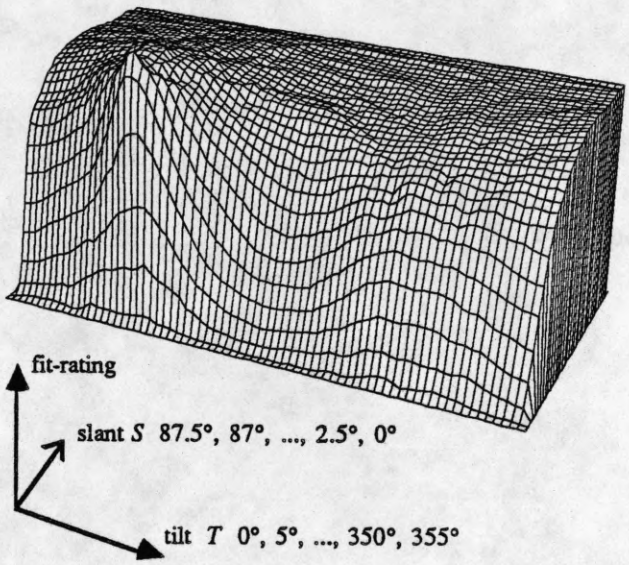
(d)

Figure 27

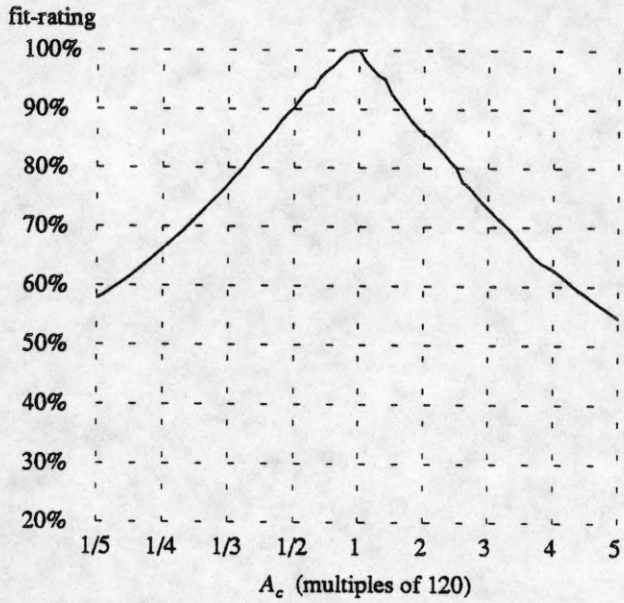
(a) Water lillies. (b) Disks corresponding to positive-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. (c) Extracted texels. These are all regions (sets of overlapping disks) having area within a factor of two of the area expected by the best planar fit (A_e 120, slant 75°, tilt 90°). The texels that fit the plane most closely are printed darkest. (d) The texels superimposed on a dark reproduction of the original.



(e)



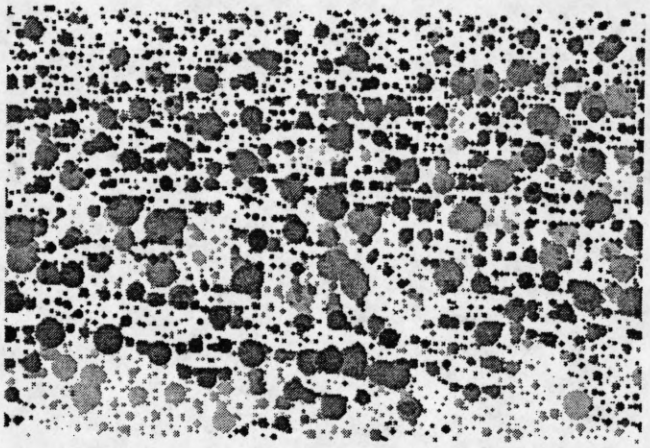
(f)



(g)

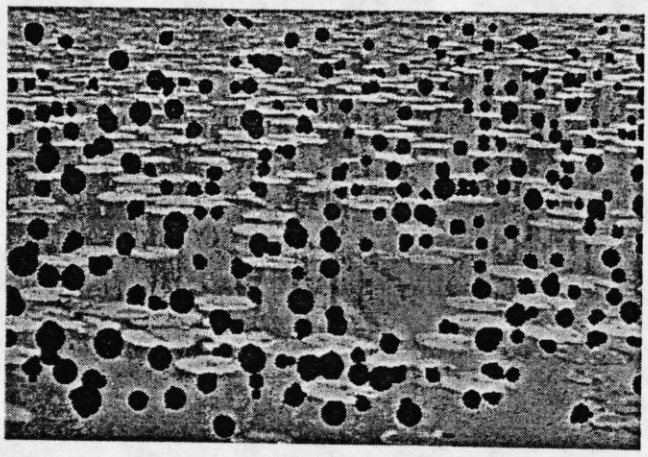
Figure 27, continued (Water lillies; positive-contrast regions)
(e) Synthetic image to illustrate the planar fit A_c 120, slant 75°, tilt 90°. (f) and (g) Ratings of various possible planar fits. In (f) slant and tilt are varied while A_c is constant at 120. In (g) A_c is varied while slant and tilt are constant at 75° and 90° respectively.

(a)



(b)

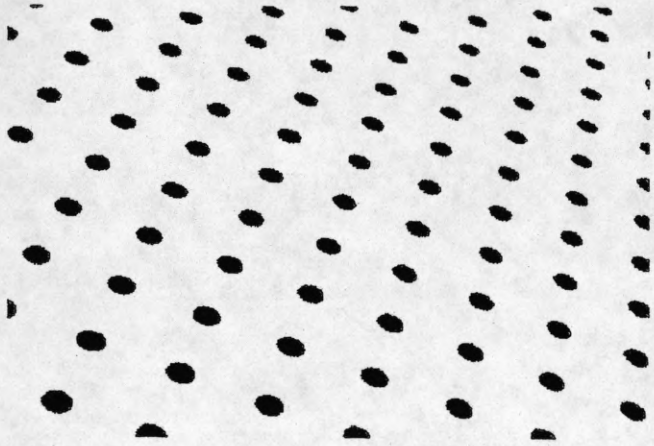
(c)



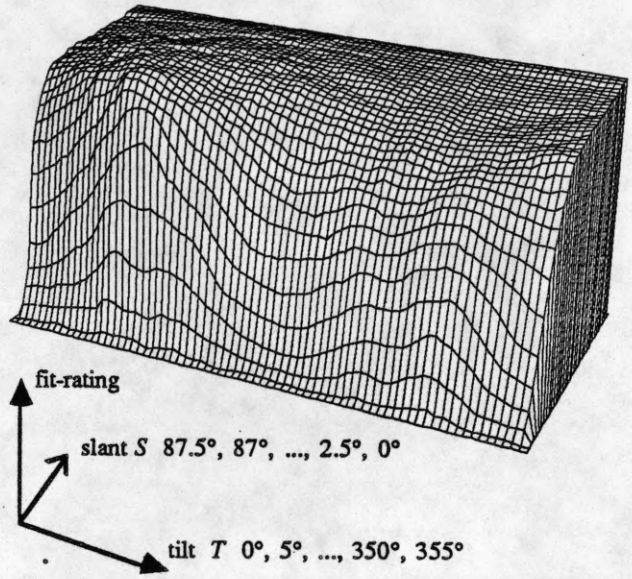
(d)

Figure 28

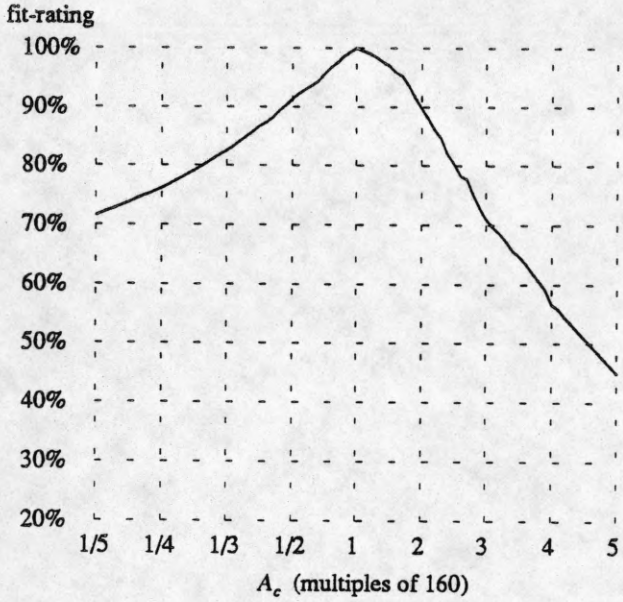
(a) Water lillies. (b) Disks corresponding to negative-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. (c) Extracted texels. These are all regions (sets of overlapping disks) having area within a factor of two of the area expected by the best planar fit (A_c 160, slant 52.5°, tilt 70°). The texels that fit the plane most closely are printed darkest. (d) The texels superimposed on a bright reproduction of the original.



(e)

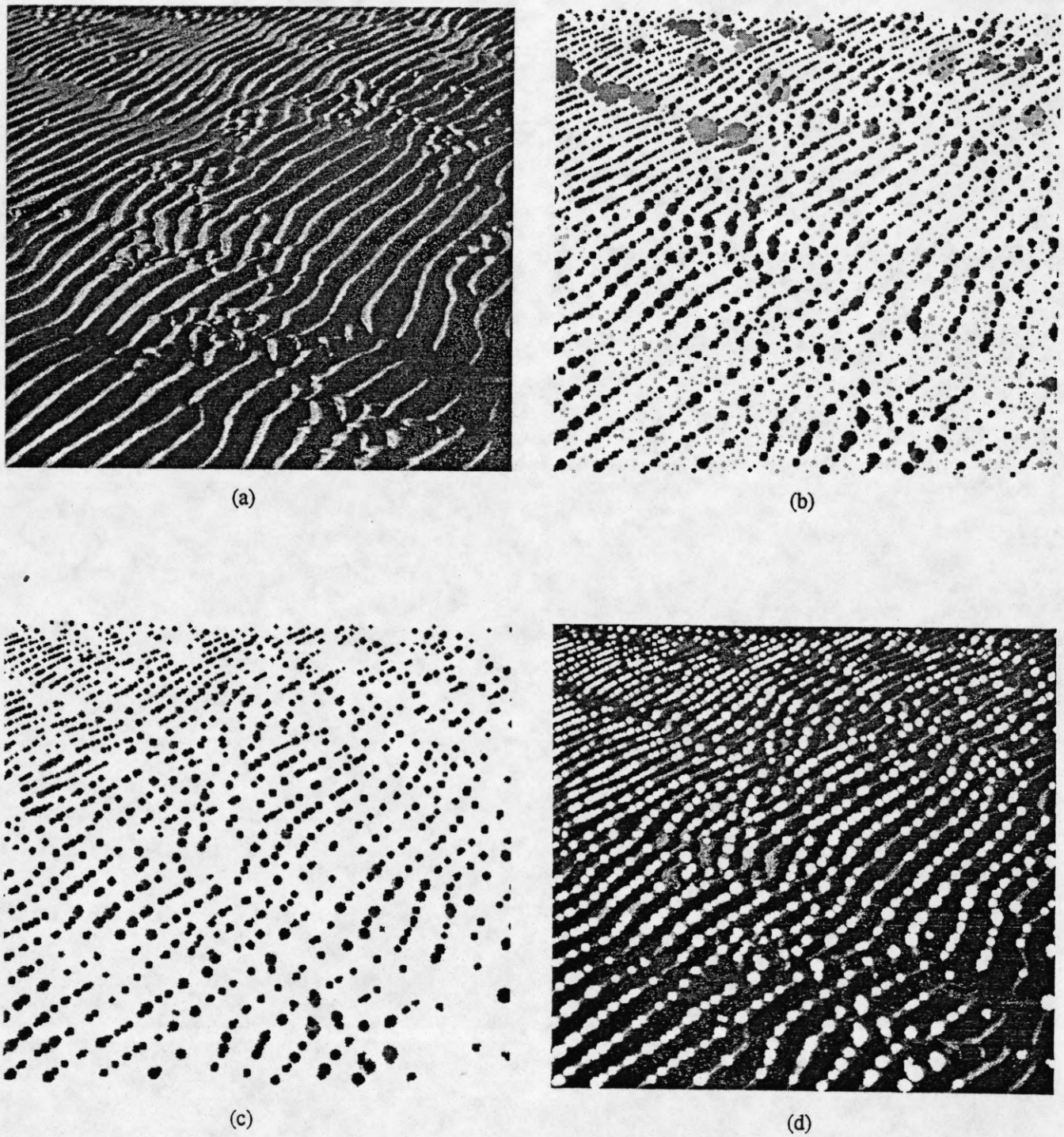


(f)



(g)

Figure 28, continued (Water lillies; negative-contrast regions)
(e) Synthetic image to illustrate the planar fit A_c 160, slant 52.5°, tilt 70°. (f) and (g) Ratings of various possible planar fits. In (f) slant and tilt are varied while A_c is constant at 160. In (g) A_c is varied while slant and tilt are constant at 52.5° and 70° respectively.

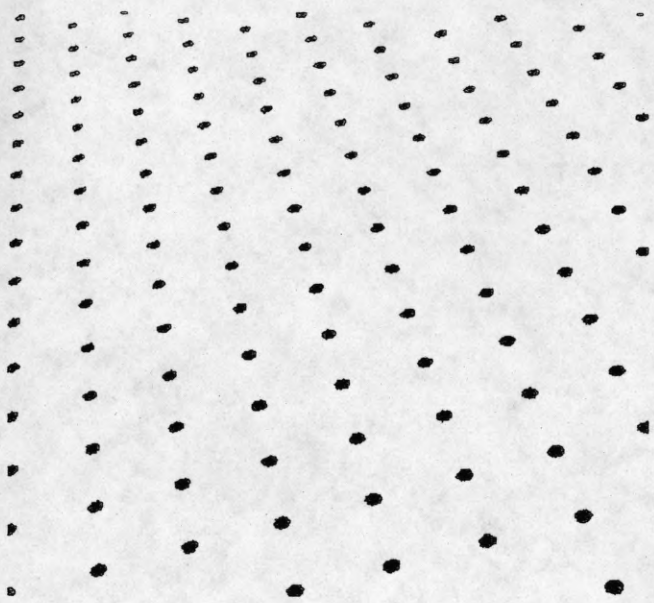


(a) (b)

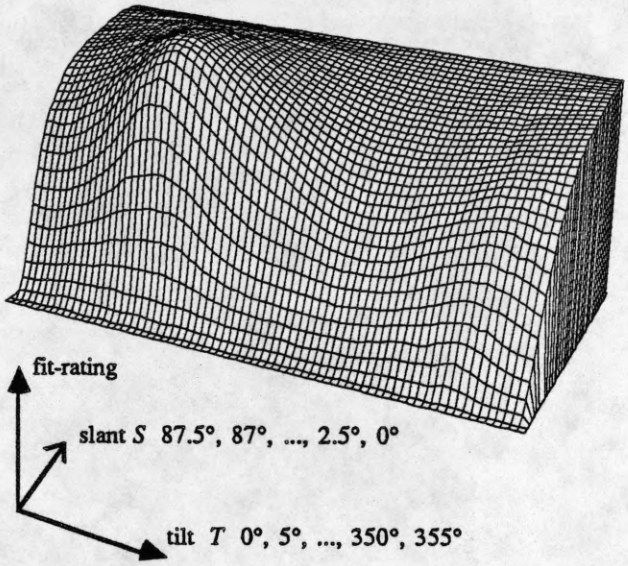
(c) (d)

Figure 29

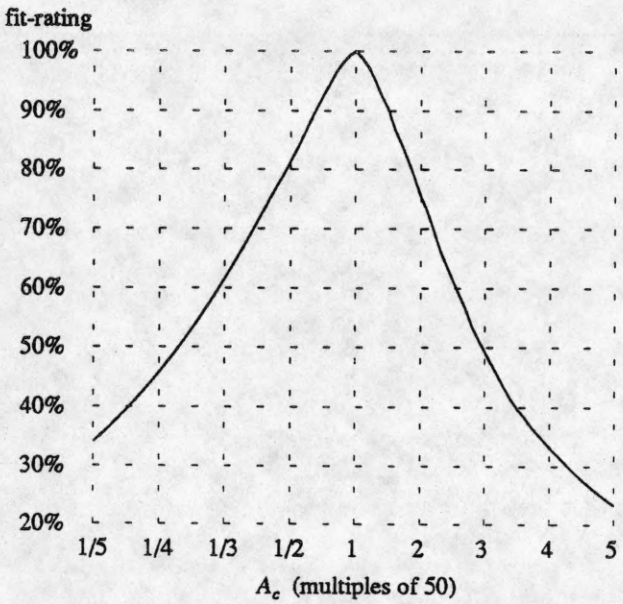
(a) Ripple marks in shallow sea. (b) Disks corresponding to positive-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. (c) Extracted texels. These are all regions (sets of overlapping disks) having area within a factor of two of the area expected by the best planar fit (A_c 50, slant 52.5°, tilt 105°). The texels that fit the plane most closely are printed darkest. (d) The texels superimposed on a dark reproduction of the original.



(e)

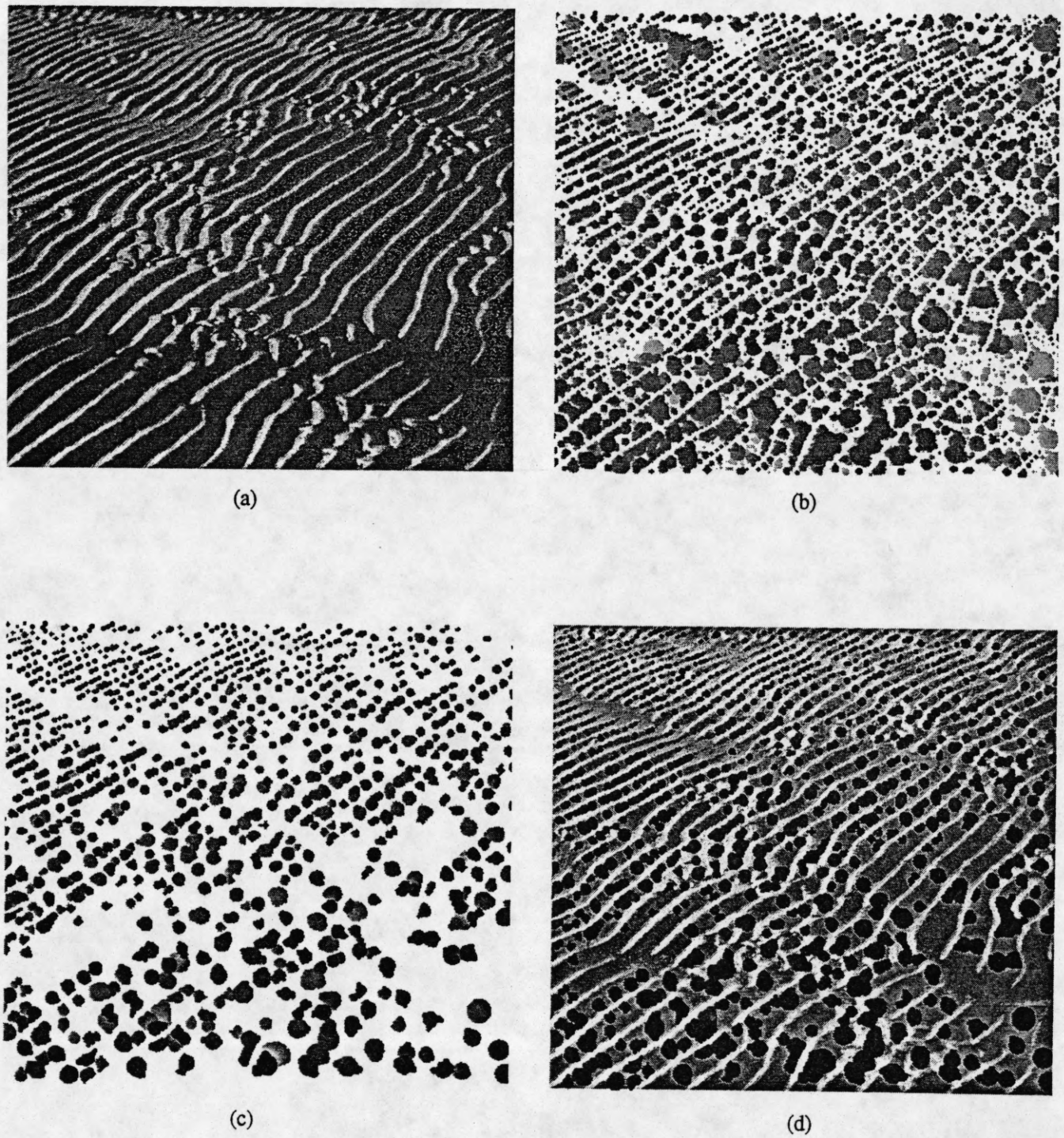


(f)



(g)

Figure 29, continued (Ripple marks in shallow sea; positive-contrast regions)
(e) Synthetic image to illustrate the planar fit A_c 50, slant 52.5°, tilt 105°. (f) and (g) Ratings of various possible planar fits. In (f) slant and tilt are varied while A_c is constant at 50. In (g) A_c is varied while slant and tilt are constant at 52.5° and 105° respectively.

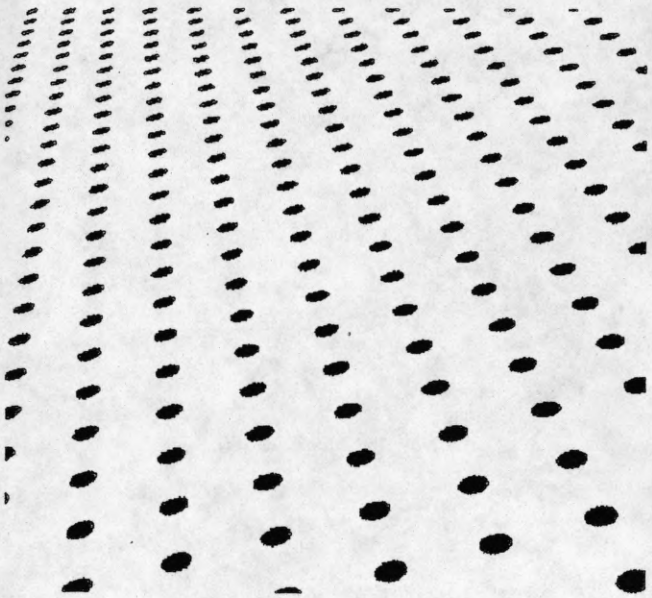


(a) (b)

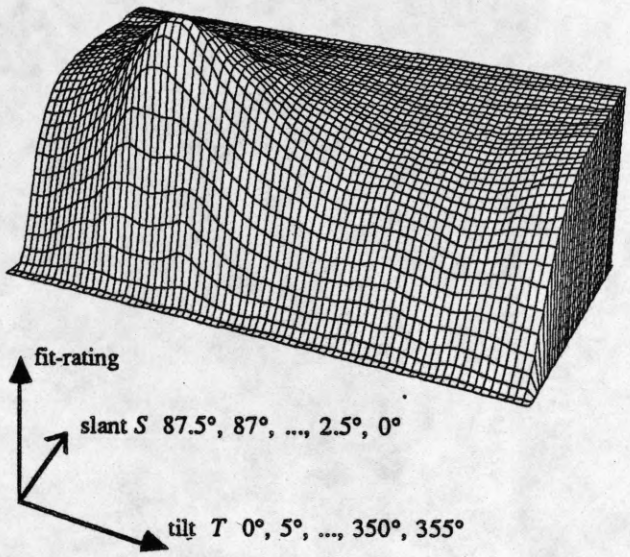
(c) (d)

Figure 30

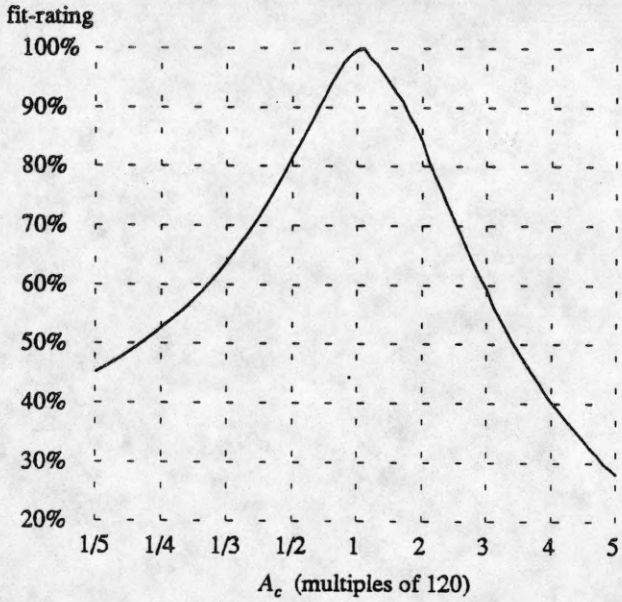
(a) Ripple marks in shallow sea. (b) Disks corresponding to negative-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. (c) Extracted texels. These are all regions (sets of overlapping disks) having area within a factor of two of the area expected by the best planar fit (A_c 120, slant 62.5°, tilt 105°). The texels that fit the plane most closely are printed darkest. (d) The texels superimposed on a bright reproduction of the original.



(e)



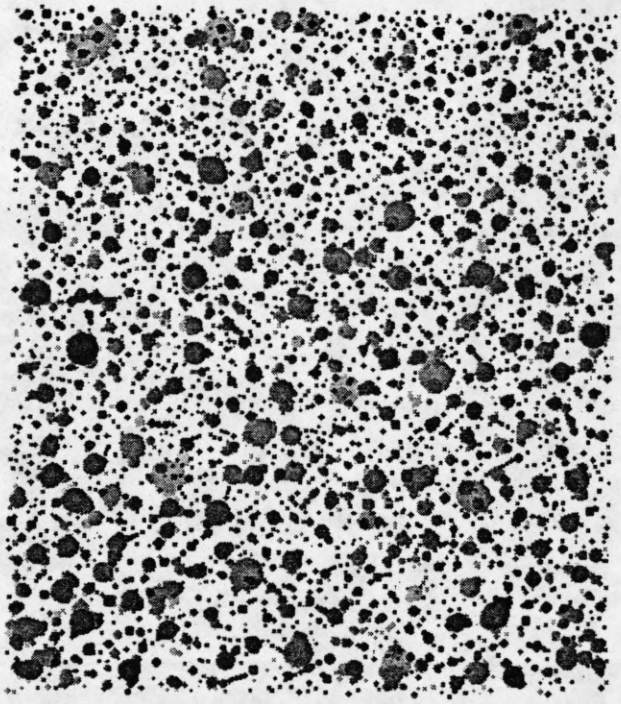
(f)



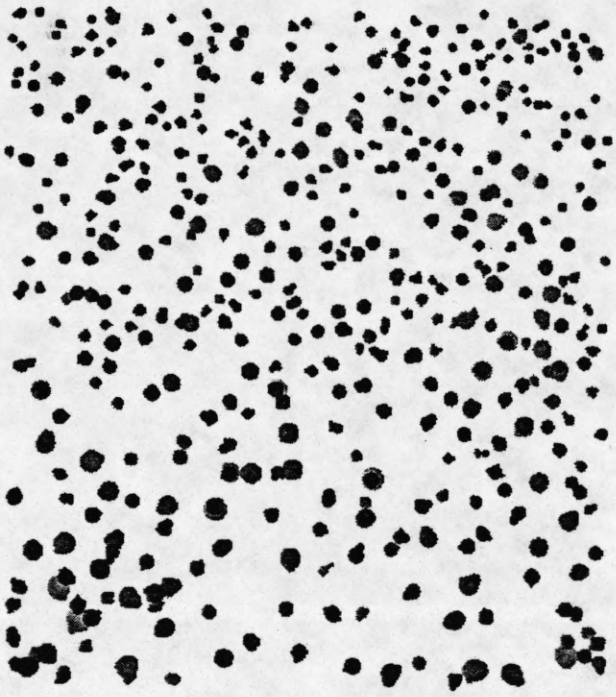
(g)

Figure 30, continued (Ripple marks in shallow sea; negative-contrast regions)
(e) Synthetic image to illustrate the planar fit A_c 120, slant 62.5°, tilt 105°. (f) and (g) Ratings of various possible planar fits. In (f) slant and tilt are varied while A_c is constant at 120. In (g) A_c is varied while slant and tilt are constant at 62.5° and 105° respectively.

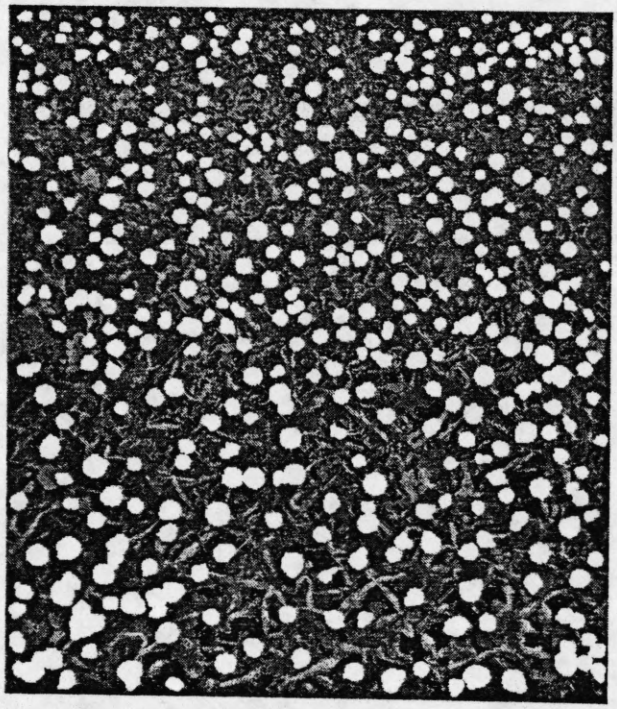
(a)



(b)



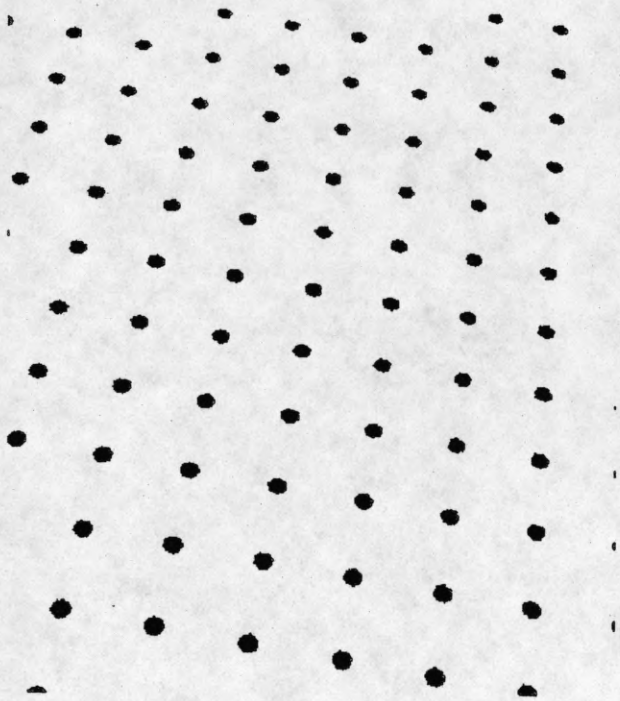
(c)



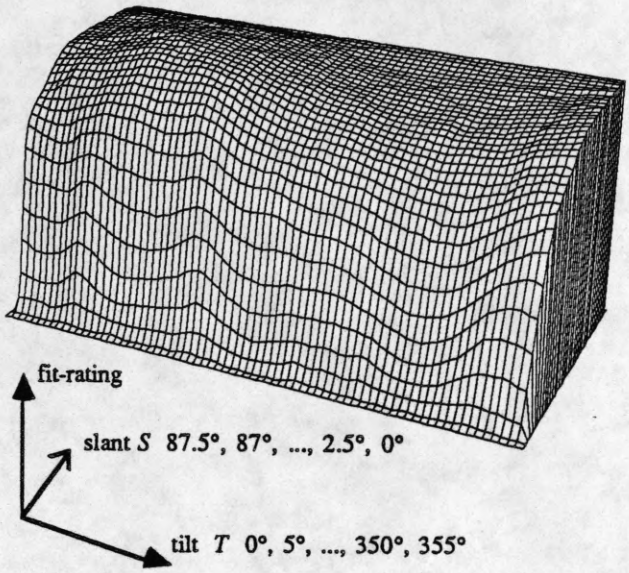
(d)

Figure 31

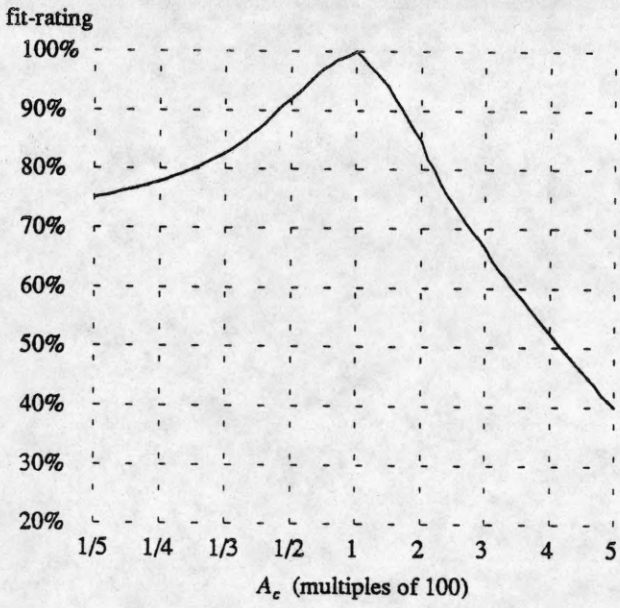
(a) Water Hyacinths. (b) Disks corresponding to positive-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. (c) Extracted texels. These are all regions (sets of overlapping disks) having area within a factor of two of the area expected by the best planar fit (A_e 100, slant 37.5°, tilt 80°). The texels that fit the plane most closely are printed darkest. (d) The texels superimposed on a dark reproduction of the original.



(e)



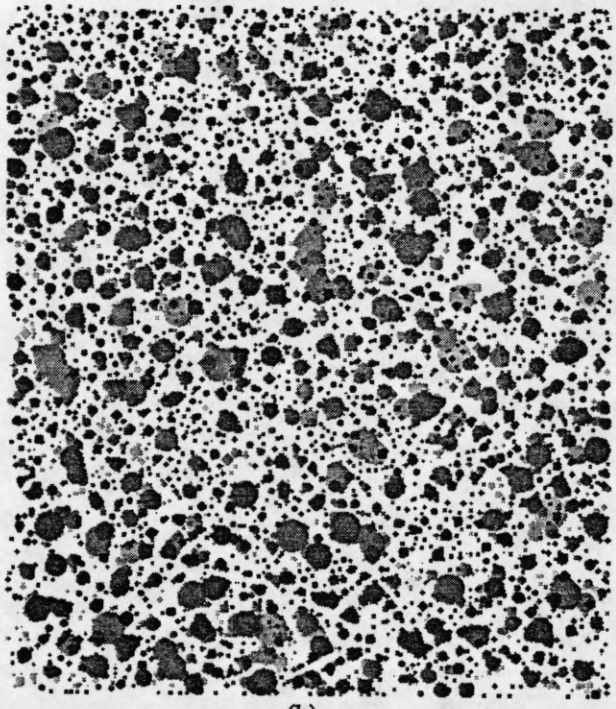
(f)



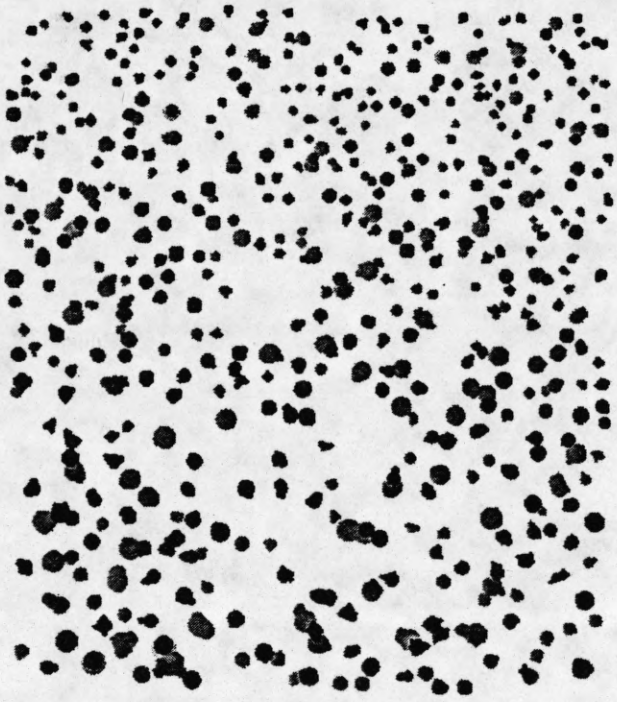
(g)

Figure 31, continued (Water Hyacinths; positive-contrast regions)
(e) Synthetic image to illustrate the planar fit A_c 100, slant 37.5°, tilt 80°. (f) and (g) Ratings of various possible planar fits. In (f) slant and tilt are varied while A_c is constant at 100. In (g) A_c is varied while slant and tilt are constant at 37.5° and 80° respectively.

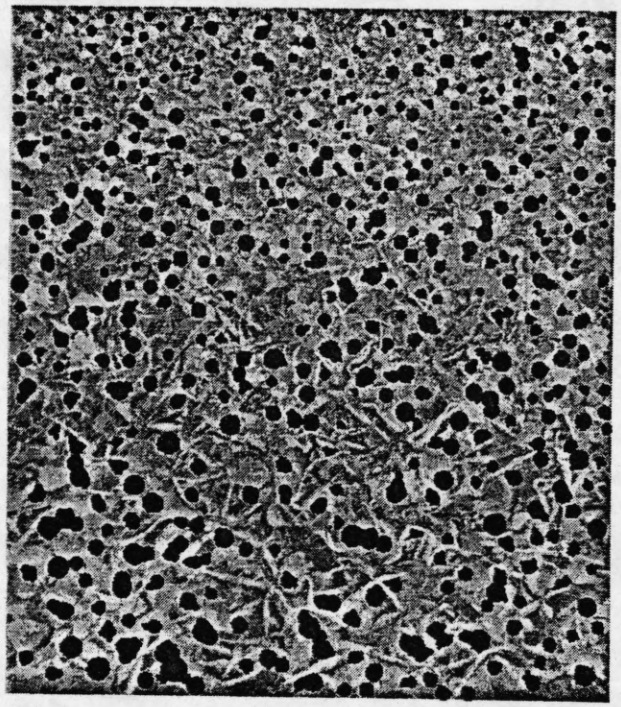
(a)



(b)



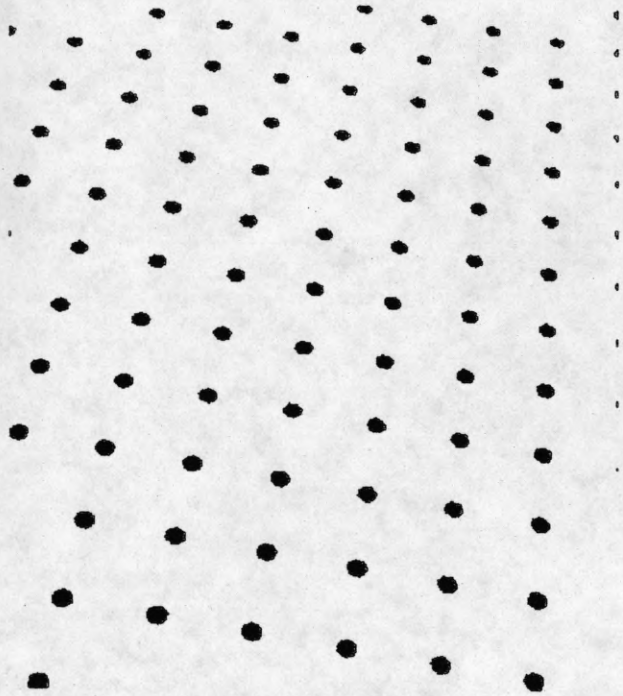
(c)



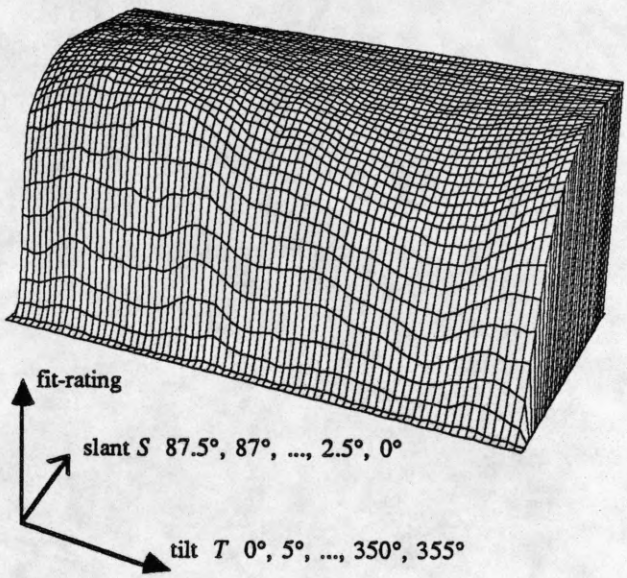
(d)

Figure 32

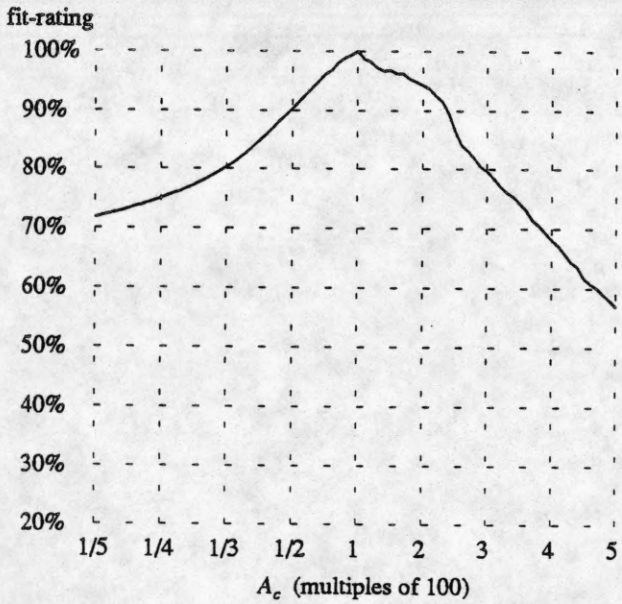
(a) Water Hyacinths. (b) Disks corresponding to negative-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. (c) Extracted texels. These are all regions (sets of overlapping disks) having area within a factor of two of the area expected by the best planar fit (A_c 100, slant 40°, tilt 80°). The texels that fit the plane most closely are printed darkest. (d) The texels superimposed on a bright reproduction of the original.



(e)

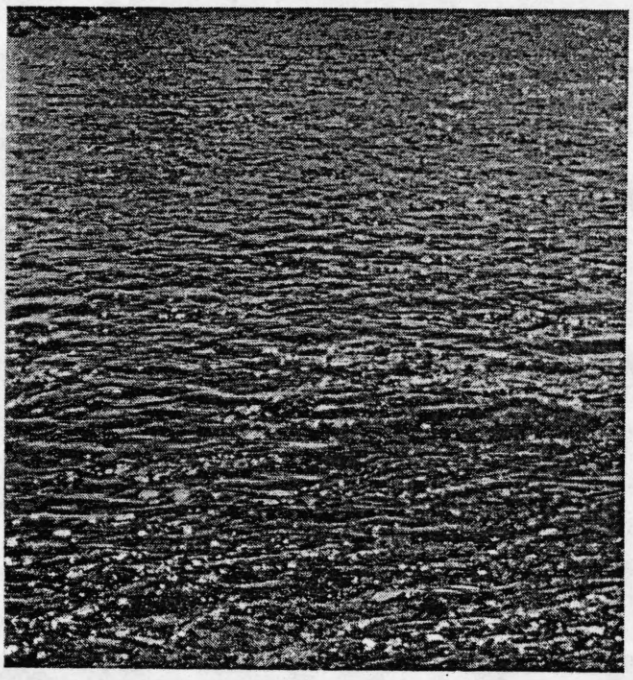


(f)

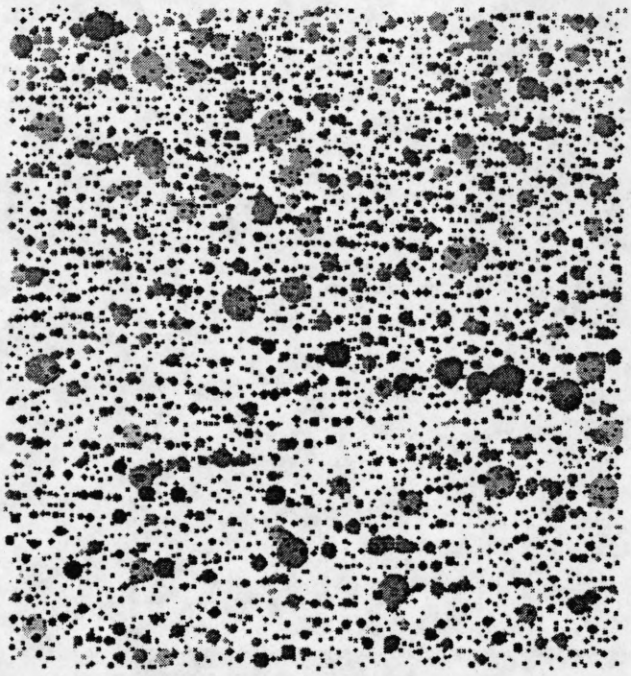


(g)

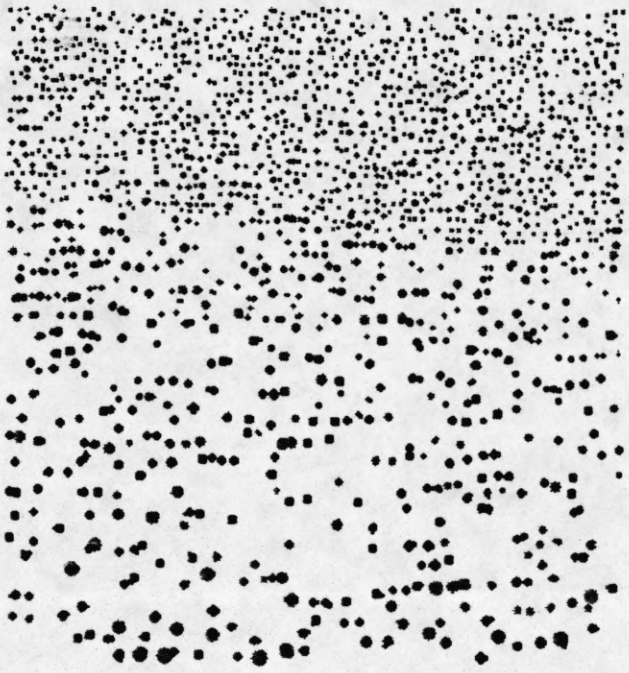
Figure 32, continued (Water Hyacinths; negative-contrast regions)
(e) Synthetic image to illustrate the planar fit A_c 100, slant 40°, tilt 80°. (f) and (g) Ratings of various possible planar fits. In (f) slant and tilt are varied while A_c is constant at 100. In (g) A_c is varied while slant and tilt are constant at 40° and 80° respectively.



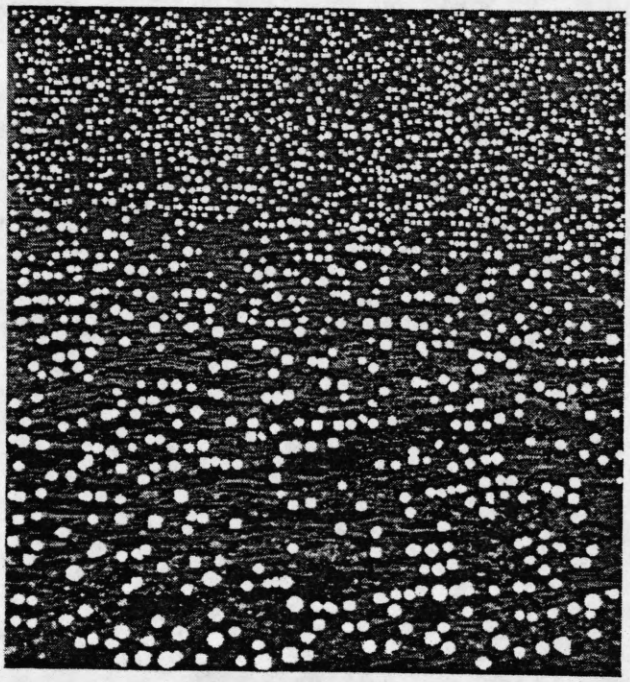
(a)



(b)



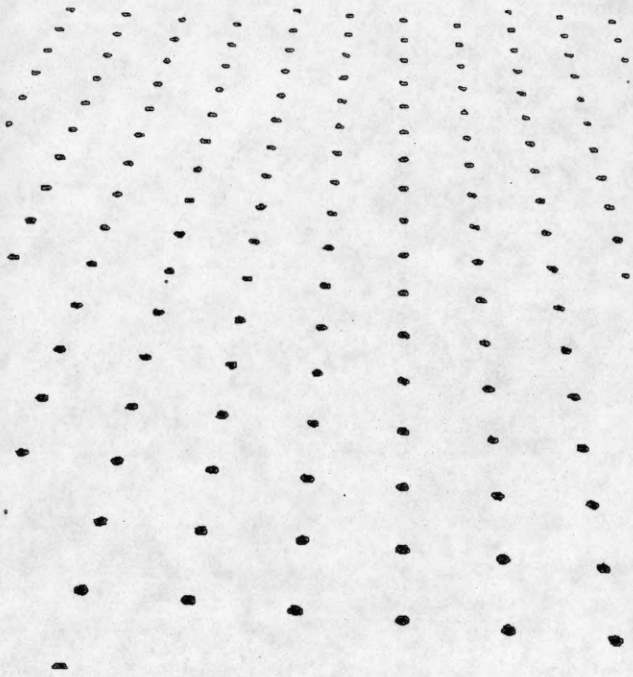
(c)



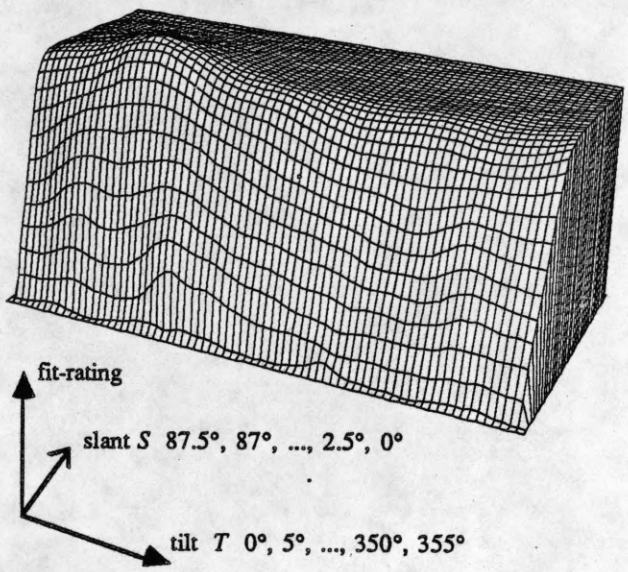
(d)

Figure 33

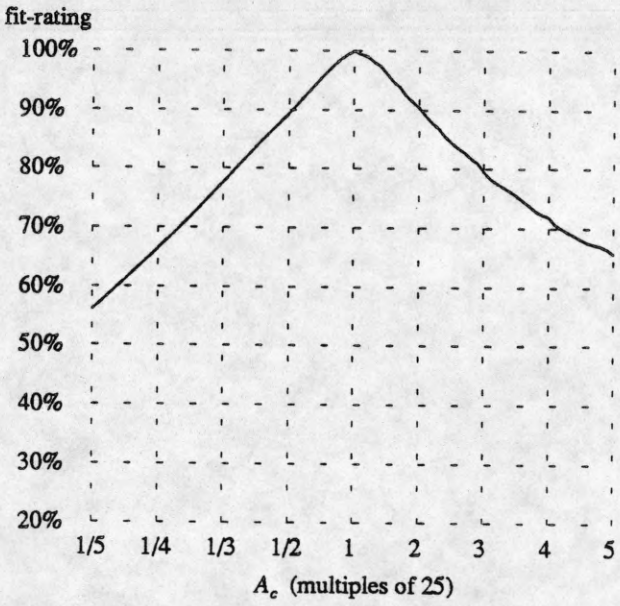
(a) The Toulumne River. (b) Disks corresponding to positive-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. (c) Extracted texels. These are all regions (sets of overlapping disks) having area within a factor of two of the area expected by the best planar fit (A_c 25, slant 57.5°, tilt 85°). The texels that fit the plane most closely are printed darkest. (d) The texels superimposed on a dark reproduction of the original.



(e)

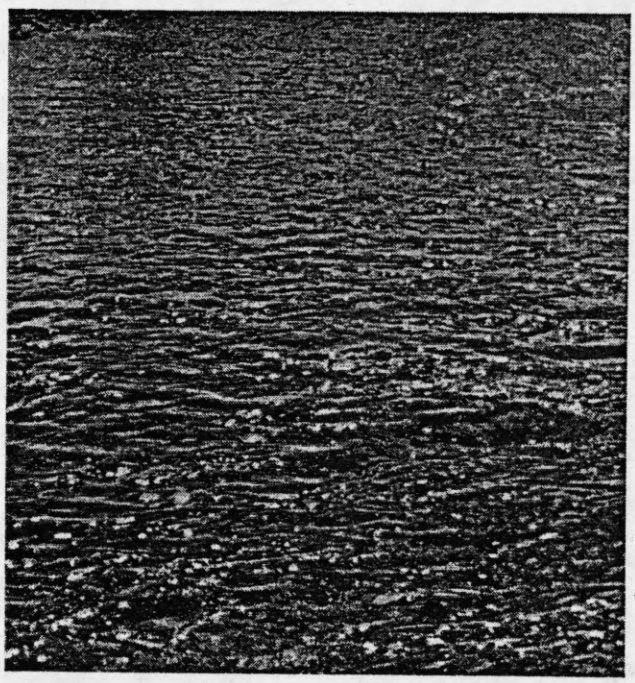


(f)

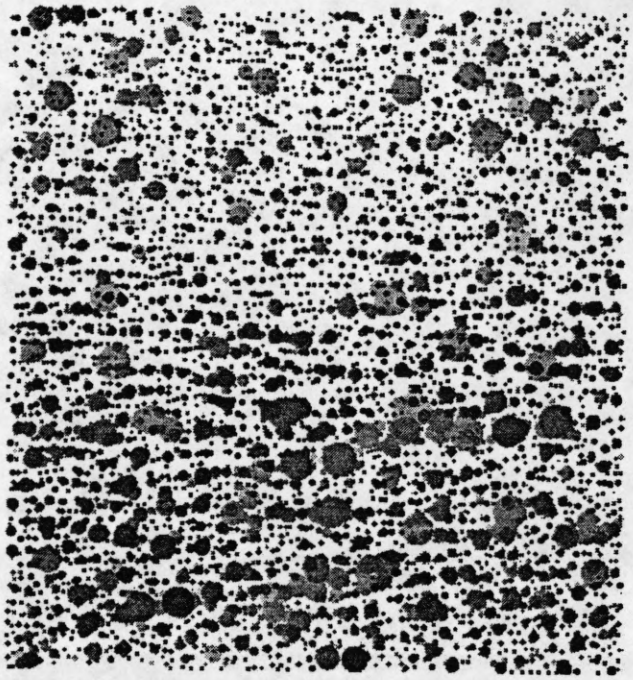


(g)

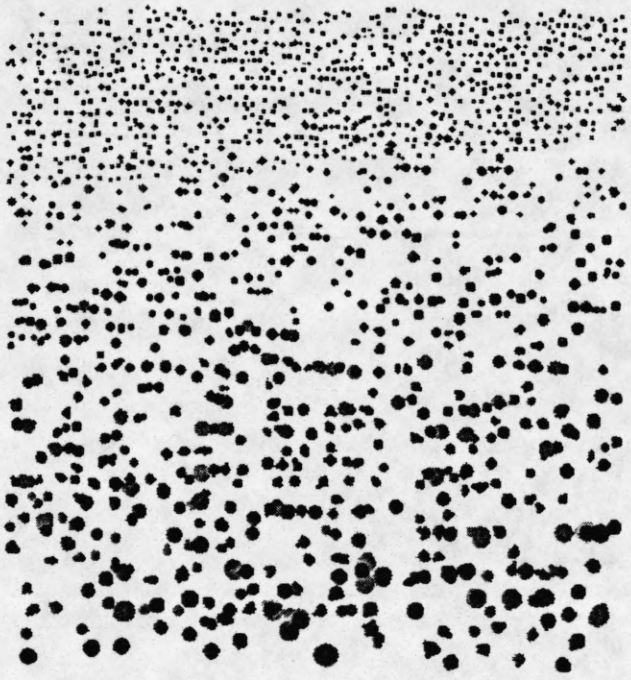
Figure 33, continued (The Toulumne River; positive-contrast regions)
(e) Synthetic image to illustrate the planar fit A_c 25, slant 57.5°, tilt 85°. (f) and (g) Ratings of various possible planar fits. In (f) slant and tilt are varied while A_c is constant at 25. In (g) A_c is varied while slant and tilt are constant at 57.5° and 85° respectively.



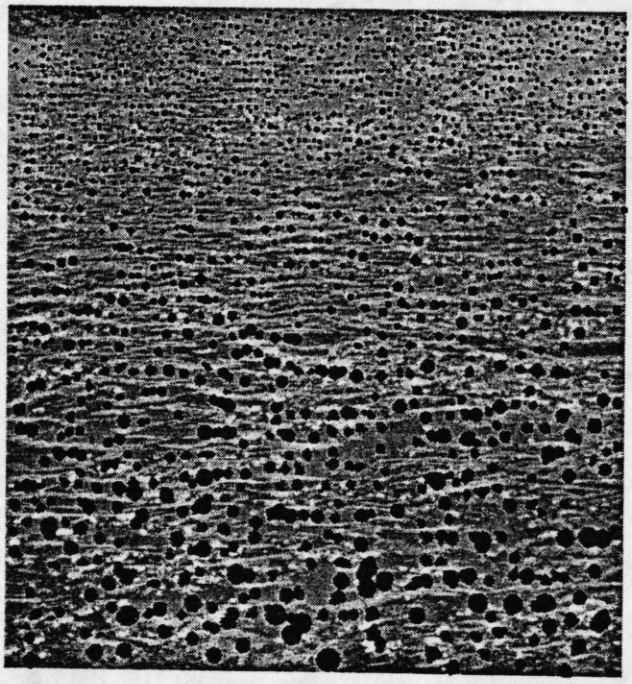
(a)



(b)



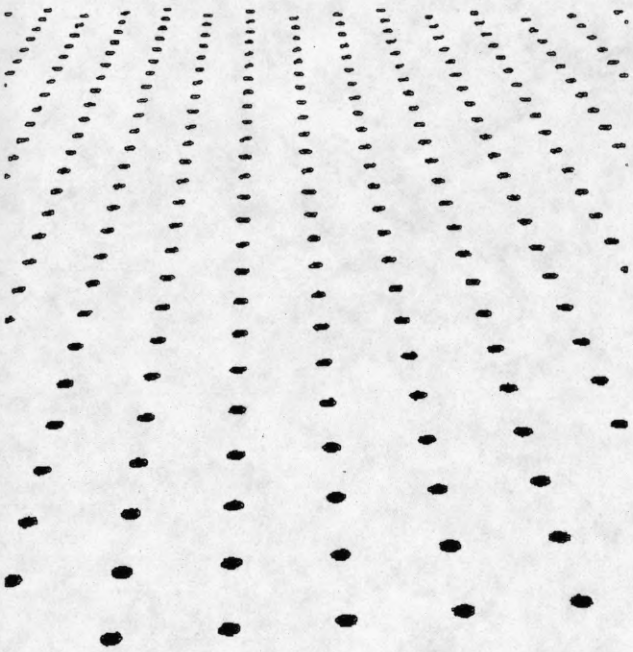
(c)



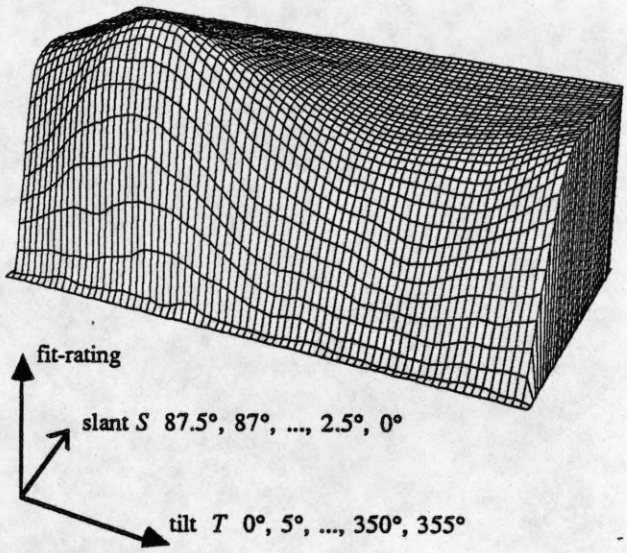
(d)

Figure 34

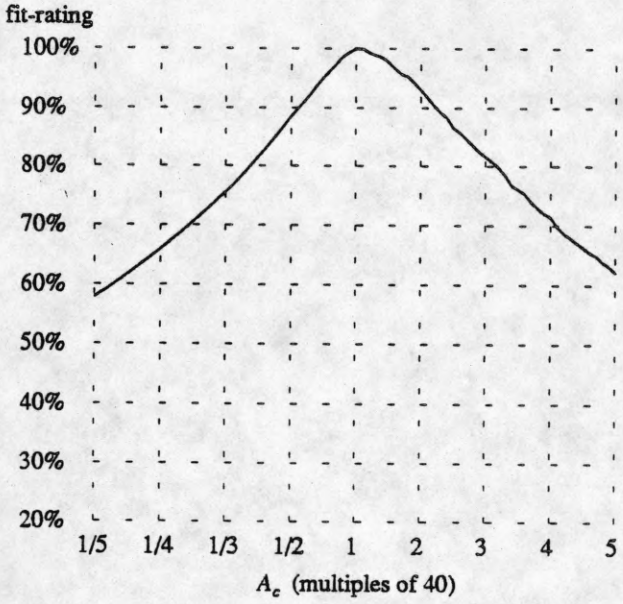
(a) The Toulumne River. (b) Disks corresponding to negative-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. (c) Extracted texels. These are all regions (sets of overlapping disks) having area within a factor of two of the area expected by the best planar fit (A_c 40, slant 65°, tilt 95°). The texels that fit the plane most closely are printed darkest. (d) The texels superimposed on a bright reproduction of the original.



(e)



(f)

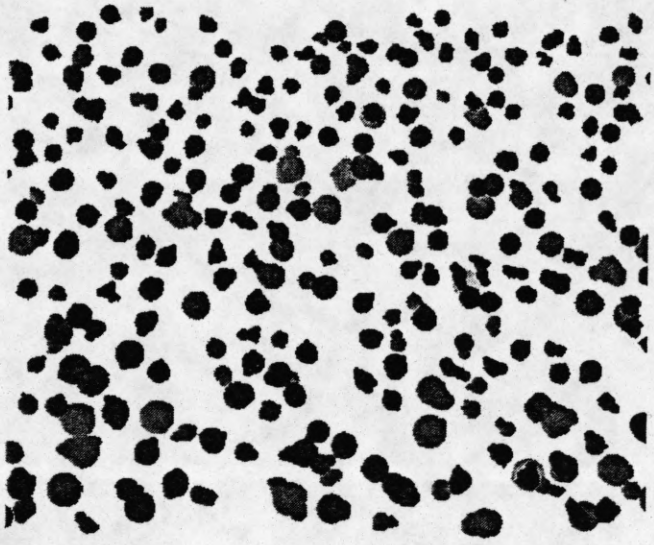


(g)

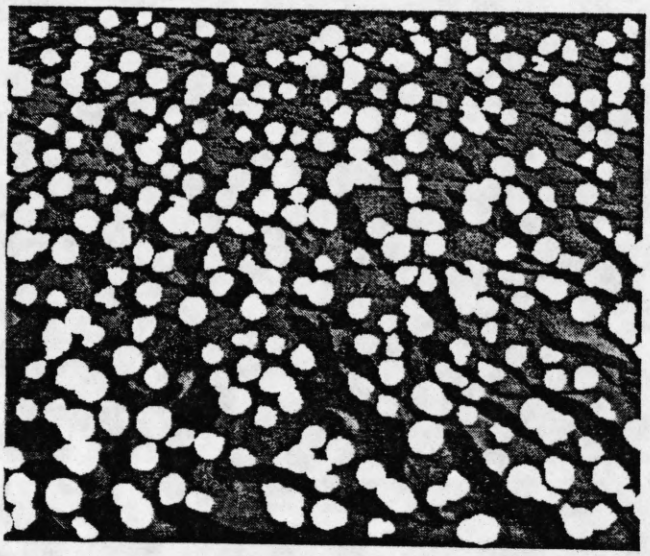
Figure 34, continued (The Toulumne River; negative-contrast regions)
(e) Synthetic image to illustrate the planar fit A_c 40, slant 65°, tilt 95°. (f) and (g) Ratings of various possible planar fits. In (f) slant and tilt are varied while A_c is constant at 40. In (g) A_c is varied while slant and tilt are constant at 65° and 95° respectively.

(a)

(b)



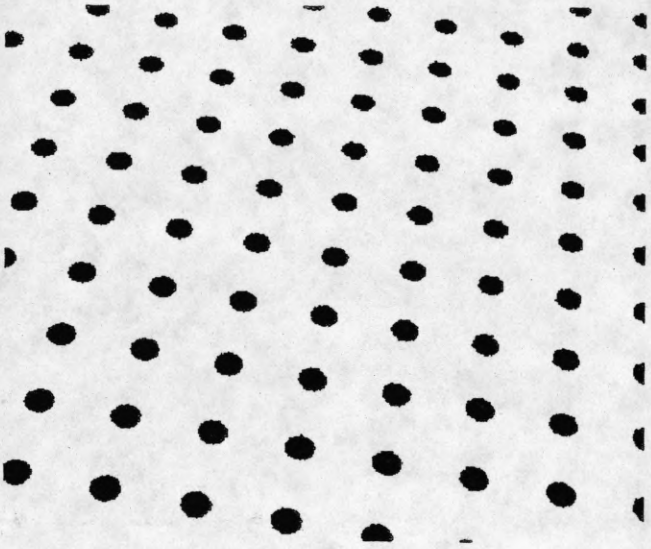
(c)



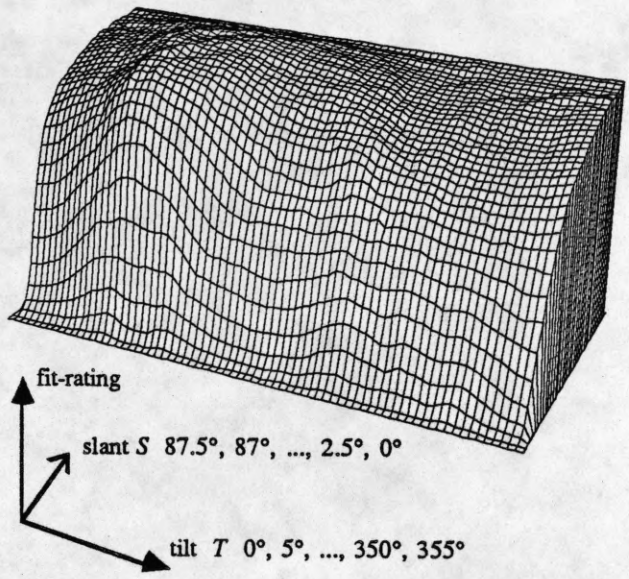
(d)

Figure 35

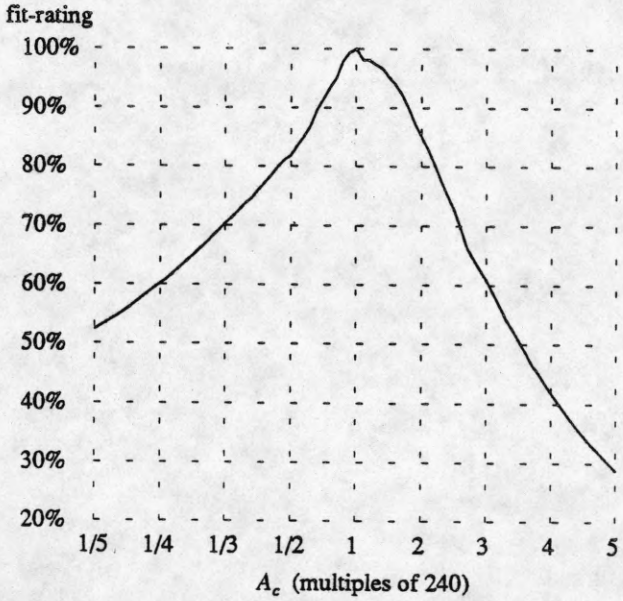
(a) Sand by the Adriatic Sea. (b) Disks corresponding to positive-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. (c) Extracted texels. These are all regions (sets of overlapping disks) having area within a factor of two of the area expected by the best planar fit (A_e 240, slant 40°, tilt 80°). The texels that fit the plane most closely are printed darkest. (d) The texels superimposed on a dark reproduction of the original.



(e)



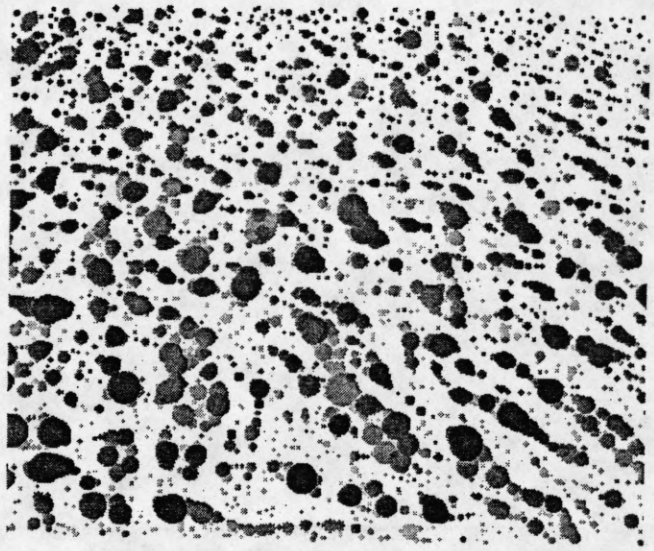
(f)



(g)

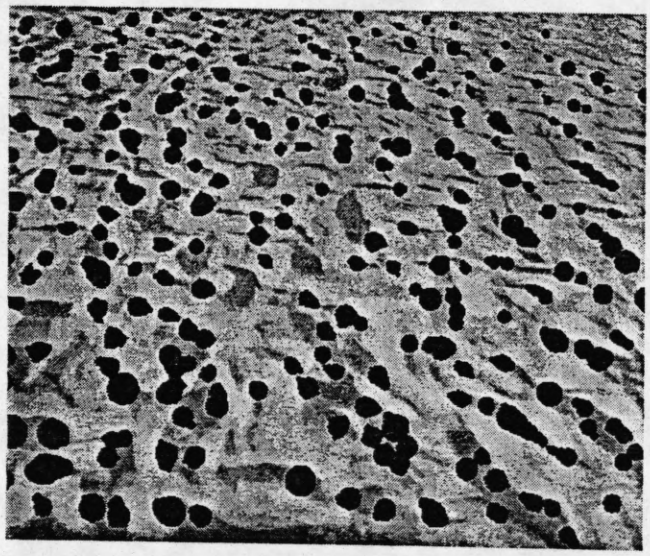
Figure 35, continued (Sand by the Adriatic Sea; positive-contrast regions)
(e) Synthetic image to illustrate the planar fit A_c 240, slant 40°, tilt 80°. (f) and (g) Ratings of various possible planar fits. In (f) slant and tilt are varied while A_c is constant at 240. In (g) A_c is varied while slant and tilt are constant at 40° and 80° respectively.

(a)



(b)

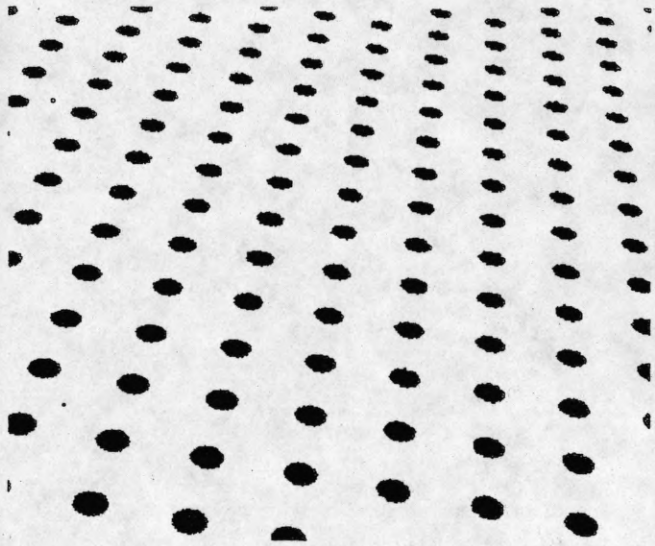
(c)



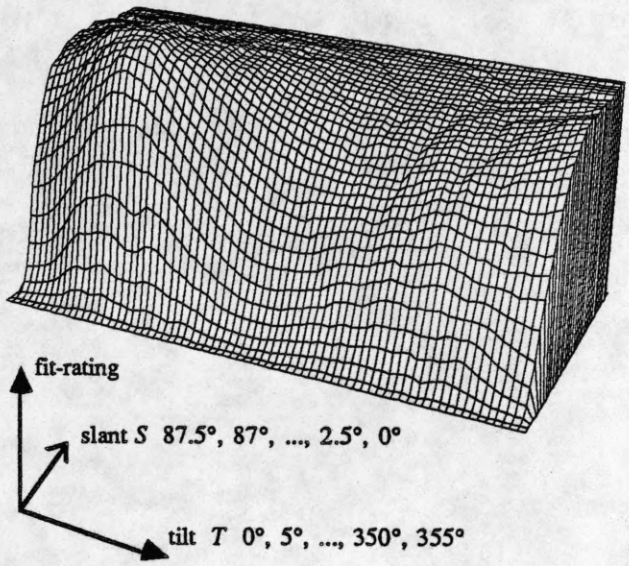
(d)

Figure 36

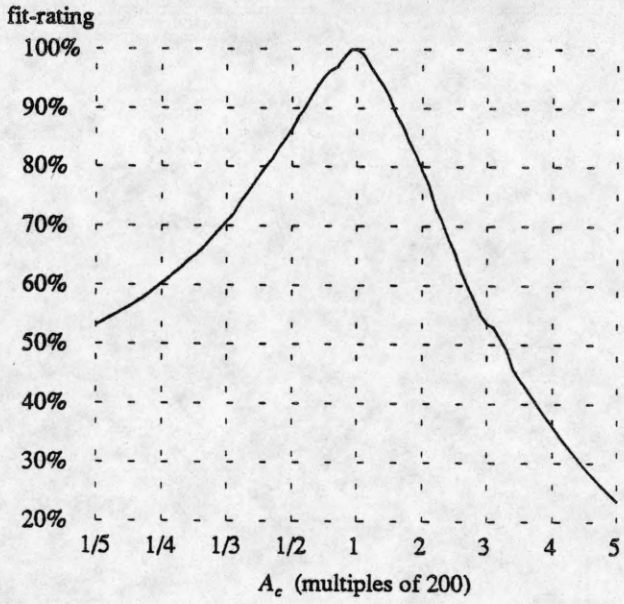
(a) Sand by the Adriatic Sea. (b) Disks corresponding to negative-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. (c) Extracted texels. These are all regions (sets of overlapping disks) having area within a factor of two of the area expected by the best planar fit (A_e 200, slant 55°, tilt 80°). The texels that fit the plane most closely are printed darkest. (d) The texels superimposed on a bright reproduction of the original.



(e)



(f)



(g)

Figure 36, continued (Sand by the Adriatic Sea; negative-contrast regions)
(e) Synthetic image to illustrate the planar fit A_c 200, slant 55°, tilt 80°. (f) and (g) Ratings of various possible planar fits. In (f) slant and tilt are varied while A_c is constant at 200. In (g) A_c is varied while slant and tilt are constant at 55° and 80° respectively.

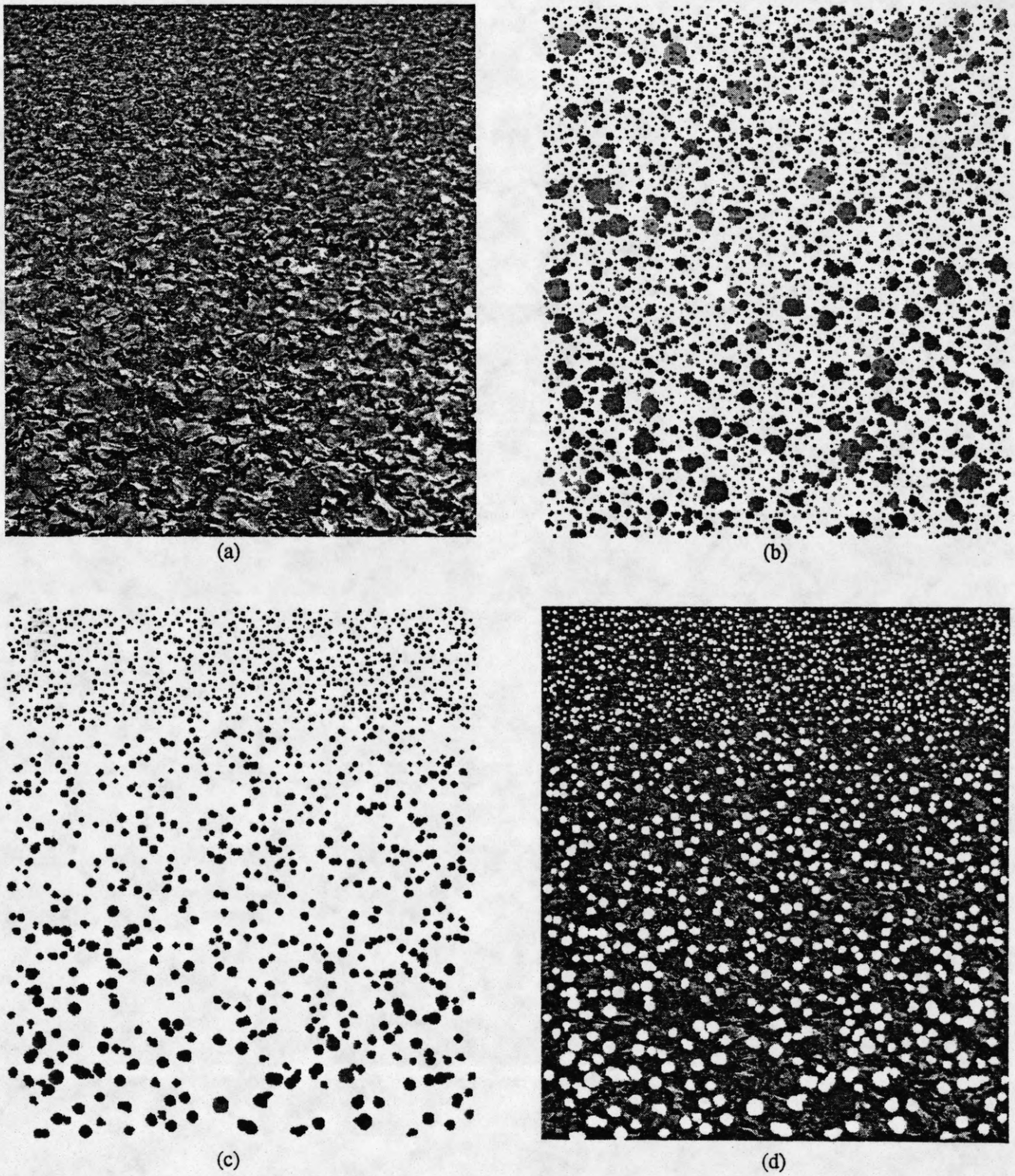
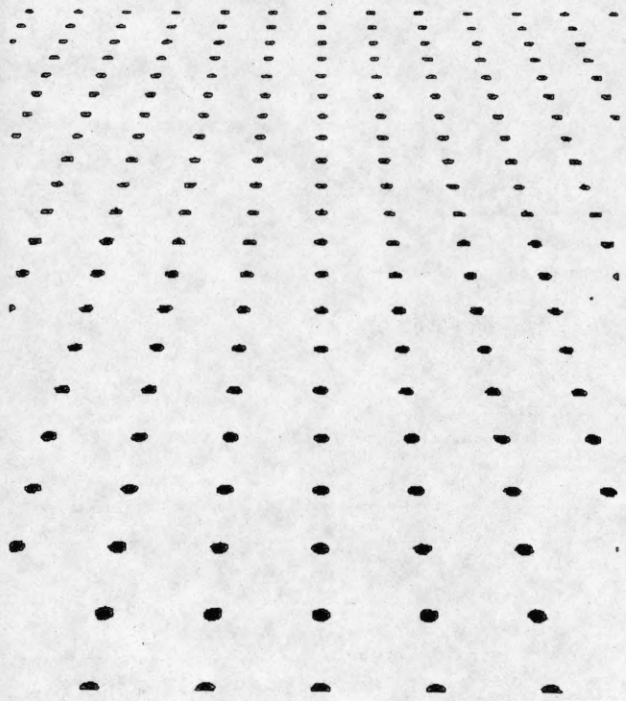
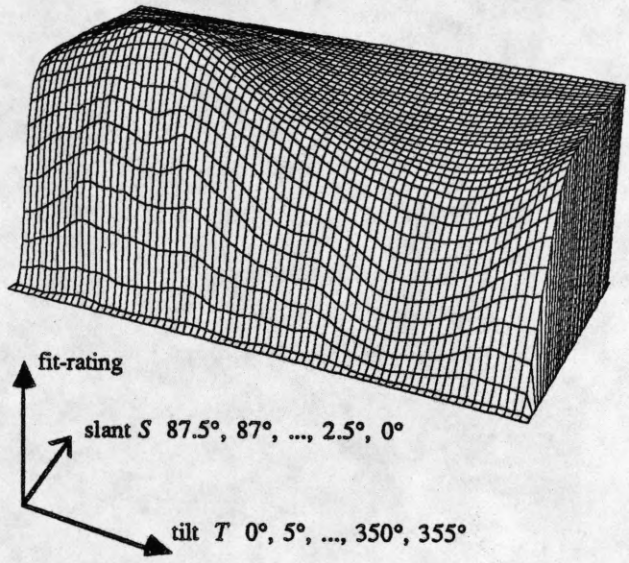


Figure 37

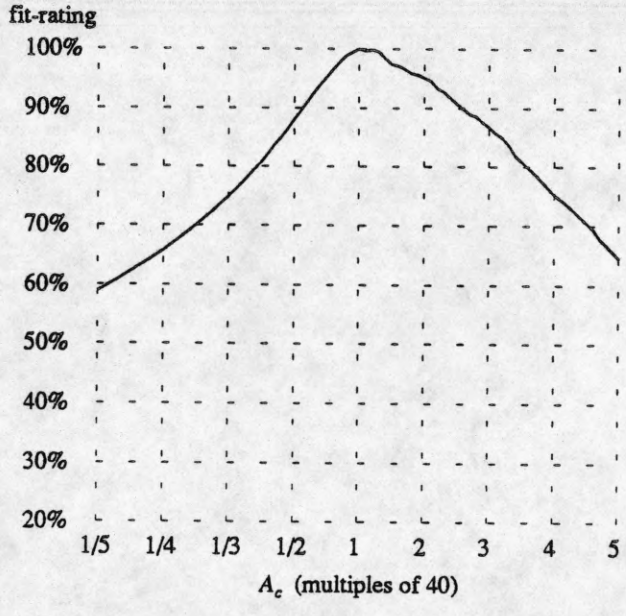
(a) Fallen leaves. (b) Disks corresponding to positive-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. (c) Extracted texels. These are all regions (sets of overlapping disks) having area within a factor of two of the area expected by the best planar fit (A_e 40, slant 60°, tilt 90°). The texels that fit the plane most closely are printed darkest. (d) The texels superimposed on a dark reproduction of the original.



(e)

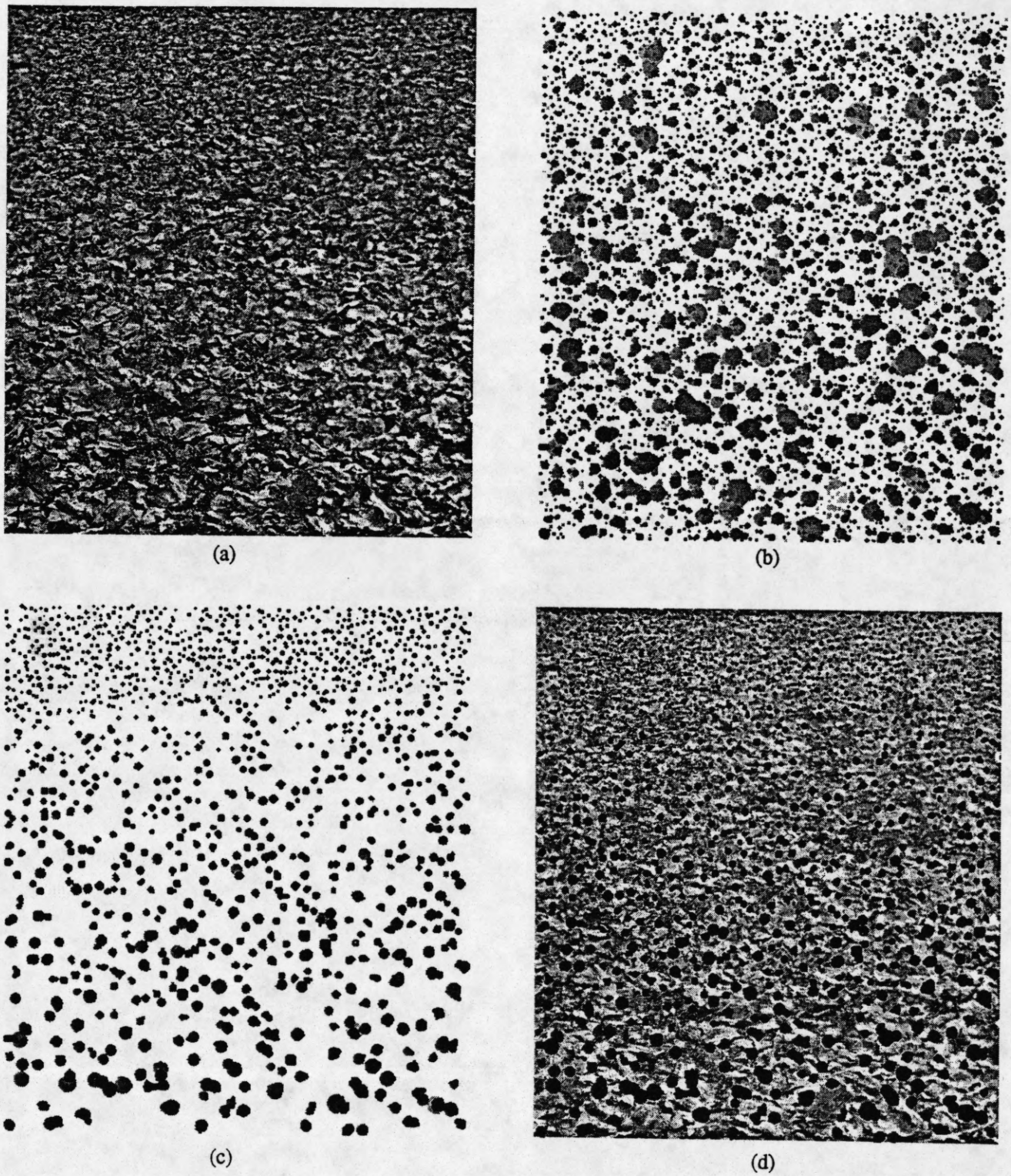


(f)



(g)

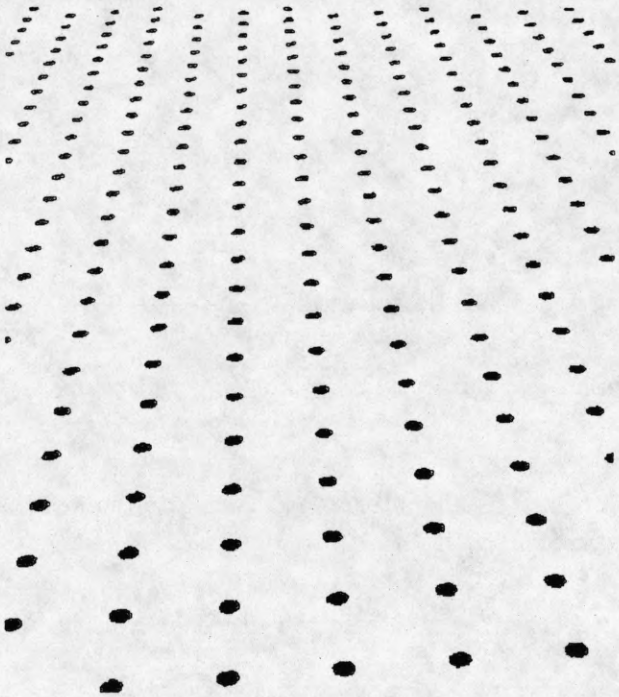
Figure 37, continued (Fallen leaves; positive-contrast regions)
(e) Synthetic image to illustrate the planar fit A_c 40, slant 60°, tilt 90°. (f) and (g) Ratings of various possible planar fits. In (f) slant and tilt are varied while A_c is constant at 40. In (g) A_c is varied while slant and tilt are constant at 60° and 90° respectively.



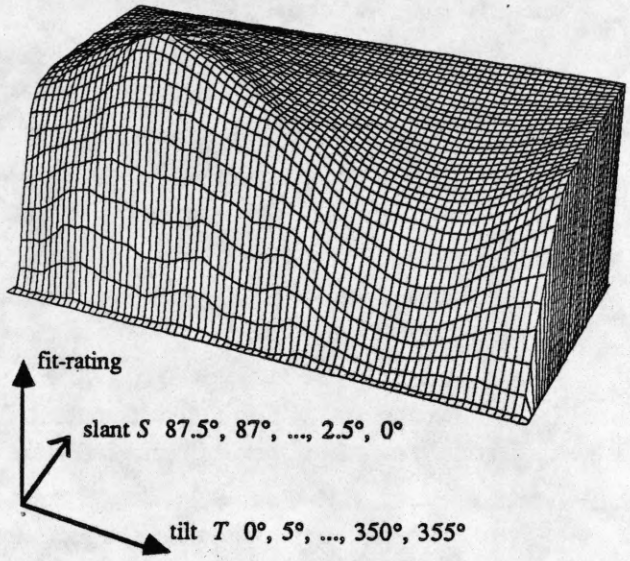
(a) (b) (c) (d)

Figure 38

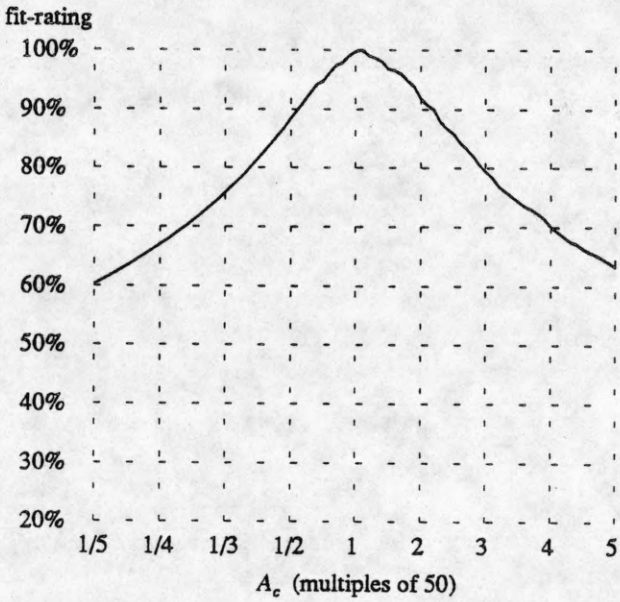
(a) Fallen leaves. (b) Disks corresponding to negative-contrast regions of relatively uniform gray level. Disks are shown with a darkness proportional to the contrast of the region. (c) Extracted texels. These are all regions (sets of overlapping disks) having area within a factor of two of the area expected by the best planar fit (A_c 50, slant 62.5°, tilt 95°). The texels that fit the plane most closely are printed darkest. (d) The texels superimposed on a bright reproduction of the original.



(e)



(f)



(g)

Figure 38, continued (Fallen leaves; negative-contrast regions)
(e) Synthetic image to illustrate the planar fit A_c 50, slant 62.5°, tilt 95°. (f) and (g) Ratings of various possible planar fits. In (f) slant and tilt are varied while A_c is constant at 50. In (g) A_c is varied while slant and tilt are constant at 62.5° and 95° respectively.

REFERENCES

Adams, A. and N. Newhall [1960].
This is the American Earth, Sierra Club, San Fransisco.

Ahuja, N. and B. Schachter [1983].
Pattern Models, Wiley, 1983.

Ahuja, N. and B. Schachter [1983b].
"Image Models", *Computing Surveys*, Vol. 13, No 4, 373-397, December 1981.

Aloimonos, J. and M. Swain [1985].
"Shape from Texture", *Proceedings of the 9th International Joint Conference on Artificial Intelligence*, 926-931, 1985.

Aloimonos, J. [1986].
"Detection of Surface Orientation from Texture I: The Case of Planes", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 584-593, 1986.

Attneave, F. and R. K. Olson [1966].
"Inferences About Visual Mechanisms from Monocular Depth Effects", *Psychonomic Science*, 4, 133-134, 1966.

Bajcsy, R. and L. Lieberman [1976].
"Texture Gradient as a Depth Cue", *Computer Graphics and Image Processing*, vol 5, 52-67, 1976.

Bourke-White, Margaret [1972].
The Photographs of Margaret Bourke-White, Edited by Sean Callahan, New York Graphic Society, Greenwich, Connecticut.

Braunstein, M. L. and J. W. Payne [1969].
"Perspective and Form Ratio as Determinants of Relative Slant Judgments", *Journal of Experimental Psychology*, 81(3), 584-590, 1969.

Brodatz, P. [1966].
Textures: A Photographic Album for Artists and Designers, Dover, New York, 1966.

Crowley, J. and A. Parker [1984].
"A Representation for Shape Based on Peaks and Ridges in the Difference of Low Pass Transform", *IEEE Pattern Analysis and Machine Intelligence*, vol 6, no 2, 156-170, March 1984.

Cutting, J. E. and R. T. Millard [1984].
"Three Gradients and the Perception of Flat and Curved Surfaces", *Journal of Experimental Psychology: General*, 113(2), 198-216, 1984.

Davis, L., L. Janos, and S. Dunn [1983].
"Efficient Recovery of Shape from Texture", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-5, No. 5, 485-492, September 1983.

Dunn, S., L. Davis, and H. Hakalahti [1984].
"Experiments in Recovering Surface Orientation from Texture", University of Maryland Computer Science Technical Report CS-TR-1399, May 1984.

Dyer, Charles R. and Azriel Rosenfeld [1976].
"Fourier Texture Features: Suppression of Aperture Effects", *IEEE Transactions on Systems, Man, and Cybernetics*, Vol 6, 703-705, October 1976.

Eriksson, S. [1964].
"Monocular Slant Perception and the Texture Gradient Concept", *Scandinavian Journal of Psychology*, Vol 5, 123-128, 1964.

Flock, H. R. [1964].
"A Possible Optical Basis for Monocular Slant Perception", *Psychological Review*, 71(5), 380-391, 1964.

Flock, H. R. [1965].
"Optical Texture and Linear Perspective as Stimuli for Slant Perception", *Psychological Review*, 72(6), 505-514, 1965.

Freeman, R. B. [1965].
"Ecological Optics and Visual Slant", *Psychological Review*, 72(6), 501-504, 1965.

Freeman, R. B. [1966a].
"Optical Texture Versus Retinal Perspective: A Reply to Flock", *Psychological Review*, 73(4), 365-371, 1966.

Freeman, R. B. [1966b].
"Effect of Size on Visual Slant", *Journal of Experimental Psychology*, 71(1), 96-103, 1966.

Gibson, J. [1950].
The Perception of the Visual World, Houghton Mifflin, Boston, 1950.

Gibson, J. [1966].
The Senses Considered as Perceptual Systems, Houghton Mifflin, Boston, 1966.

Grimson, W. E. L. and E. C. Hildreth [1985].
"Comments on "Digital Step Edges from Zero Crossing of Second Directional Derivatives", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol PAMI-7, No. 1, 121-129, January 1985.

Gruber, H. E. and W. C. Clark [1956].
"Perception of Slanted Surfaces", *Perceptual and Motor Skills*, 6, 97-106, 1956.

Gullers, Karl W. and B. Strandell [1977].
Linnaeus, Gullers International, Sweden.

Haralick, R. [1979].
"Statistical and Structural Approaches to Texture", *Proceedings of the IEEE*, Vol 67, No 5, May 1979, 786-804.

Ikeuchi, K. [1980].
"Shape from Regular Patterns (An Example of Constraint Propagation in Vision)", MIT A.I. Memo 567, March 1980.

Kanatani, K. [1984].
"Detection of Surface Orientation and Motion from Texture by a Stereological Technique", *Artificial Intelligence*, 23, 213-237, 1984.

Kanatani, K. and T. Chou [1986].
"Shape from Texture: General Principle", *Proceedings Computer Vision and Pattern Recognition 86*, Miami, 578-583, June 1986.

Kender, J. [1978].

"Shape from Texture: A Brief Overview and a New Aggregation Transform", *Proceedings of the DARPA Image Understanding Workshop*, 78-84, November 1978.

Kender, J. [1979].

"Shape from Texture: A Computational Paradigm", *Proceedings of the DARPA Image Understanding Workshop*, 134-138, April 1979.

Kender, J. [1980a].

Shape from Texture, Ph. D. Thesis, Carnegie-Mellon University Computer Science Department, CMU-CS-81-102, November 1980.

Kender, J. and T. Kanade [1980b].

"Mapping Image Properties into Shape Constraints: Skewed Symmetry, Affine Transformable Patterns, and the Shape-from-Texture Paradigm", *Proceedings of the National Conference on Artificial Intelligence*, American Association for Artificial Intelligence, 4-6, 1980.

Kender, J. [1983].

"Surface Constraints from Linear Extents", *Proceedings of the American Association for Artificial Intelligence Conference*, 187-190, 1983.

Landscape [1984].

Landscape Photography, edited by D. Earnest and M. Bulzone, American Photographic Book Publishing, New York

Life [1984].

LIFE: The Second Decade 1946-1955, Little, Brown and Company, Boston.

Marr, D. and E. Hildreth [1980].

"Theory of Edge Detection", *Proceedings of the Royal Society of London*, B 207, 187-217, 1980

Marr, D. [1982].

Vision, Freeman, San Francisco, 1982.

Muerle, J. [1970].

"Some Thoughts on Texture Discrimination by Computer", *Picture Processing and Psychopictorics*, Lipkin and Rosenfeld eds, New York: Academic Press, 371-379, 1970.

Nakatani, H., S. Kimura, O. Saito and T. Kitahashi [1980].

"Extraction of Vanishing Point and its Application to Scene Analysis Based on Image Sequence", *Proceedings of the International Conference on Pattern Recognition*, 370-372, 1980.

Nevatia, R. and K. R. Babu [1980].

"Linear Feature Extraction and Description", *Computer Graphics and Image Processing*, 13, 257-269, 1980.

Ohta, Y., K. Maenobu and T. Sakai [1981].

"Obtaining Surface Orientation from Texels under Perspective Projection", *Proceedings of the International Joint Conference on Artificial Intelligence*, 746-751, 1981.

Phillips, R. J. [1970].

"Stationary Visual Texture and the Estimation of Slant Angle", *Quarterly Journal of Psychology*, 22, 389-397, 1970.

Rosenfeld, A. [1975].
"A Note on Automatic Detection of Texture Gradients", *IEEE Transactions on Computers*, vol C-24, 988-991, October 1975.

Rosinski, R. R. [1974].
"On the Ambiguity of Visual Stimulation: A Reply to Eriksson", *Perception and Psychophysics*, 16(2), 259-263, 1974.

Rosinski, R. and N. Levine [1976].
"Texture Gradient Effectiveness in the Perception of Surface Slant", *Journal of Experimental Child Psychology*, 22, 261-271, 1976.

Silverman, J. [1983].
For the World to See: the Life of Margaret Bourke-White, Viking Press, New York.

Stevens, K.A. [1981].
"The Information Content of Texture Gradients", *Biological Cybernetics*, vol 42, 95-105, 1981.

Stevens, K.A. [1983a].
"Slant-Tilt: The Visual Encoding of Surface Orientation", *Biological Cybernetics*, vol 46, 183-195, 1983.

Stevens, K.A. [1983b].
"Surface Tilt (The Direction of Slant): A Neglected Psychophysical Variable", *Perception and Psychophysics*, 33(3), 241-250, 1983.

Strache, Wolf [1956].
Forms and Patterns in Nature, Pantheon Books, New York.

Thomas, Bill [1976].
The Swamp, W. W. Norton & Company, Inc., New York.

Van Gool, L., P. Dewaele, and A. Oosterlinck [1985].
"SURVEY: Texture Analysis Anno 1983", *Computer Vision, Graphics, and Image Processing*, 29, 336-357, March 1985.

Vickers, D. [1971].
"Perceptual economy and the impression of visual depth", *Perception and Psychophysics*, 10(1), 23-27, 1971.

Witkin, A.P. [1981].
"Recovering Surface Shape and Orientation from Texture", *Artificial Intelligence*, vol 17, 17-45, 1981.

Witkin, A. P. [1983].
"Scale Space Filtering", *Eighth International Joint Conference on Artificial Intelligence*, Karlsruhe, West Germany, 1019-1022, August 1983.

Zucker, Steven W., Azriel Rosenfeld, and Larry S. Davis [1975].
"Picture Segmentation by Texture Discrimination", *IEEE Transactions on Computers*, C-24, No. 12, 1228-1233, December 1975,